VISUALIZATION OF CORRELATIONS BETWEEN PLACES OF MUSIC LISTENING AND ACOUSTIC FEATURES

Narumi Kuroko
Ochanomizu University
Tokyo, Japan

Hayato Ohya
RecoChoku Co.,Ltd.
Tokyo, Japan

Takayuki Itoh
Ochanomizu University
Tokyo, Japan

Nicolas Grossmann
TU Wien
Vienna, Austria

Hsiang-Yun Wu
TU Wien
Vienna, Austria

***Abstract*—Users often choose songs with respect to special situations and environments. We designed and developed a music recommendation method inspired by this fact. This method selects songs based on the distribution of acoustic features of the songs listened by a user at particular places that have higher ordinariness for the user. It is important to verify the relationship between the places where the songs are listened to and the acoustic features in this. Hence, we conducted the visualization to explore potential correlations between geographic locations and the music features of single users. In this paper, we designed an interactive visualization tool methods and results for the analysis of the relationship between the places and the acoustic features while listening to the songs.**

***Index Terms*—Music Information Processing, Visualization, Lifelog, Ordinariness, Machine Learning**

I. INTRODUCTION

Dissemination of new technologies on music recommendation is expected due to changes in the environment such as streaming services. Research on music recommendation has already been diversified, and a variety of music recommendation methods according to various targets and situations have been proposed. For example, a system for recommending background music (BGM) while working [1], music recommendation systems for jogging [2] and driving [3] have been proposed. Meanwhile, we hypothesize that users will enjoy music recommendations even more if the system offers a situation- and location-dependent song recommendations. This is derived from a study by North et al. [4] on the value of music in everyday life. According to this survey, calm music is preferred in quiet places such as homes and libraries, while music that feels encouragement is preferred in a noisy place such as a shopping center or sports gym.

Based on the above example, we are developing a music recommendation system that takes into account ”ordinariness” [5] that quantifies usualness and specialness values calculated from the range of daily activities of the user. This system recommends the songs based on the hypothesis that the user wants to listen to different songs between places with high degrees of ordinariness and other places. Specifically, we have a hypothesis that people listen to the songs that have certain acoustic feature values more often at places with a high degree of ordinariness. On the contrary, it is difficult to specify what kinds of acoustic features are preferable at places with a low degree of ordinariness. Instead, it is often preferable to recommend songs related to the events, landmarks, and impressions of the places.

In this work, we hypothesize that people tend to listen to similar songs (i.e., songs with similar acoustic features) in different places they are visiting on a regular basis. We refer to these places as ”places with a high degree of ordinariness.” We had a trial of visualization to verify the existence of correlations between each place and the songs listened in the place where we stay for a long time on a daily basis. This visualization verifies our hypothesis on the correlation between places and acoustic features of songs. We expect more satisfactory music recommendation systems can be developed by considering this correlation. This paper introduces the visualization method and results that verify the correlations between the locations and the acoustic feature values.

II. RELATED WORK

A. Influence of listening environment on music selection

Greenwald et al. [6] reported that emotions and attitudes of people often affect their choices. We assume emotions and attitudes are also related to the song selection. We investigated how the environment in which song is listened affects the user’s mood and song selection based on the study by Reynolds et al. [7]. The results shown in Figures 1 suggest that the environment in which the listener is placed equally affects the listener’s mood and song selection. We concluded that mood is an important factor in song selection, indicating that the environment strongly influences mood and song selection.

III. VISUALIZATION OF LOCATIONS AND ACOUSTIC FEATURE VALUES

In this section, we present the processes of our visualization that represents the relationship between locations and music feature values. Section III-A describes the recording of the user’s location information. Section III-B describes the ordinariness calculation for recognizing places where a user stays for a long time on a daily basis. Section III-C describes the calculation of acoustic feature values applying

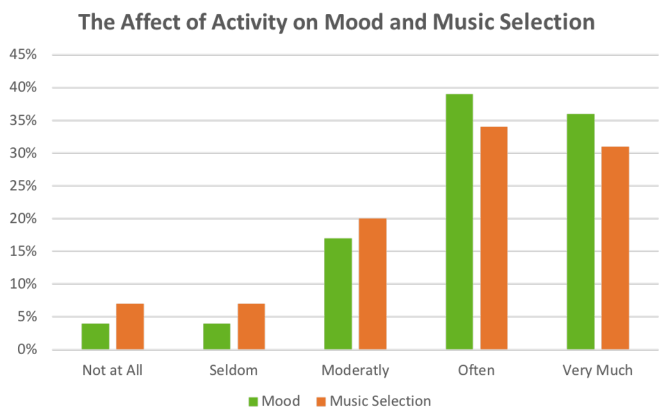

Fig. 1. Represents the level of affect that activity has on a listener's mood and music selection.

a machine learning technique. Finally, Section III-D describes the visualization method.

A. Location Information

This process records the history of location information of the users in order to calculate the ordinariness of each location. Our implementation record the locations (latitude and longitude) and the time with regular intervals using a smartphone application.

B. Ordinariness

The process divides the regions that positions are recorded into a set of ten meters squares for each location and calculates the density of the instances from all the location information recorded in advance. The calculated density is regarded as the degree of ordinariness in this study. Figure 2 show to explain visually in this process.

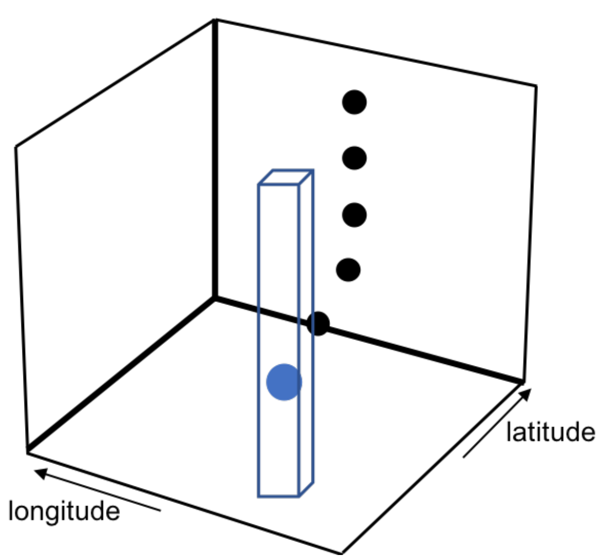

Fig. 2. How to calculate the degree of ordinariness.

Figure 3 shows an example of applying this process. The points plotted in the figure are the instances of user location information recorded every ten minutes by the process described in Section III-A. The latitude and longitude are assigned to x- and y-axes, respectively, and the z-axis corresponds to date and time. We calculate the ordinariness of the places where the user stayed occasionally for a long time smaller by taking into account the date as well as time. Meanwhile, we assume the ordinariness larger where the recorded location information is dense.

The process calculates the ordinariness applying the equation 1.

$$P = (x/M){\cdot}m \tag{1}$$

Here, let the ordinariness be P, the number of records of the location information in the range of ten meters square be x, and the maximum value of the number of location information included in the range be M. An arbitrary parameter m is adjusted so that the ordinariness value at a daily place always becomes 1. We set $m = 2.46$ in our experiments. On the other hand, the ordinary value of a particular place is zero if the staying time at the location was significantly short or the user has not been there.

The degree of ordinariness is represented by the colormap painted in orange and green in Figure 3: orangish points depict the higher degrees of ordinariness.

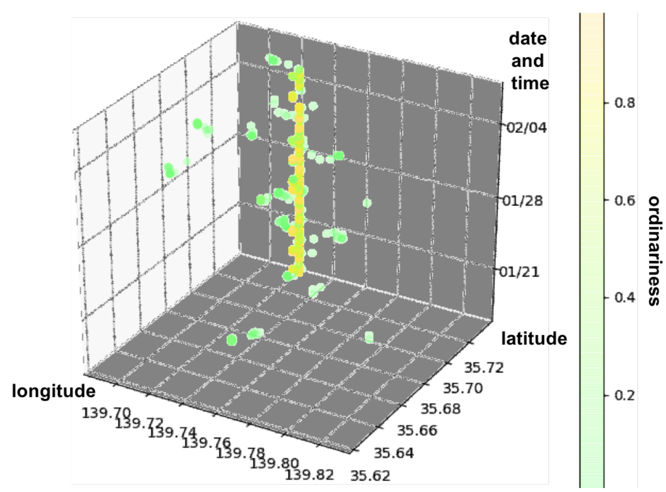

Fig. 3. Life log analysis results.

C. Acoustic Feature Values

The process calculates the acoustic feature values of each song and treats them as multi-dimensional values. We apply acoustic feature analysis tools [1] developed by the music information retrieval team of Vienna University of Technology at present. Table I shows an overview and the number of dimensions of the acoustic feature values that we have adopted.

[1] http://ifs.tuwien.ac.at/mir/audiofeatureextraction.html

TABLE I
LIST OF ACOUSTIC FEATURE VALUES UNDER ANALYSIS

Rhythm Patterns	Rhythm periodicity of the whole song Number of dimensions : 1440
Statistical Spectrum Descriptor	Beat and volume of specific range Number of dimensions : 168
Rhythm Histogram	Rhythm fluctuation of the whole song Number of dimensions : 60

D. Visualization

Finally, we apply the following two types of visualization to the ordinariness values (see Section III-B) and acoustic feature values (see Section III-C).

Figure 4 shows the visualization of the degree of ordinariness and the place where the songs are listened to. The vertical and horizontal axes correspond to latitude and longitude, respectively, in this visualization. The circular plot represents the ordinariness of the user, where orangish colors depict the places where the degree of ordinariness is larger, while greenish colors depict the places where the degree of ordinariness is smaller. Black triangular plots represent the places where the songs were listened to. This implementation displays textual information when a user points a triangular plot by a mouse-over operation. The displayed information includes latitude and longitude of the place, the degree of the ordinariness, and the title of the song listened at the place.

Figure 5 shows the visualization of the similarity of acoustic feature values of the prepared songs. Songs are arranged in a two-dimensional space by applying a dimension reduction method to the acoustic feature values. Users can select one of the following dimension reduction methods: Principal Component Analysis (PCA) [8], t-distributed Stochastic Neighbor Embedding (tSNE) [9], Multi Dimensional Scaling (MDS) [10], and Uniform Manifold Approximation and Projection (UMAP) [11], in the current implementation.

The implementation has a linked view mechanism. When a user selects a particular song or a place of interest, the visualization displays or emphasizes the details including the acoustic feature values of the song and the ordinariness of the place. This mechanism enables users to recognize the correlation between the locations where the songs were listened to and the acoustic feature values, and more detailed information recorded in the logs.

IV. EXAMPLE

This section shows visualization examples with logs of a single user. In this example, location information and music listening history logs were recorded for about one month. As a result, 1630 location information data and 98 music viewing history were recorded. The visualization results using those logs are introduced in the following sections.

A. Distribution of Ordinariness

As shown in Figure 4, the visualization window visualizes the distribution of ordinariness values and the places where the songs were listened to. A particular region where has a higher ordinariness value was enlarged and displayed in this figure in all location information. It depicts the user had two places where she had daily routines. We found these places were around the home and the university campus of the user by searching for the names of the places from their latitude and longitude values. We assumed that two places, home and university campus, are the places where the user has a high degree of ordinariness. Then, we analyzed the characteristics of the songs the user preferred to listen to at the places.

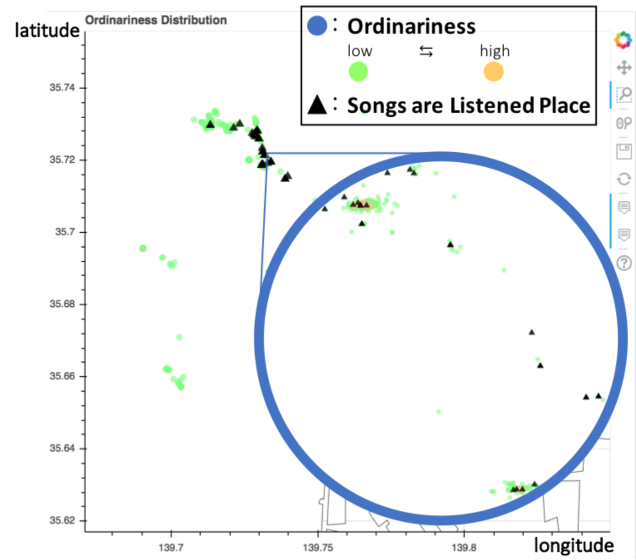

Fig. 4. Distribution of ordinariness and enlarged view.

B. Similarity of Acoustic Features

Based on the results described in Section IV-A, the songs in the playlist of the user were classified as shown in Table II. Figure 5 shows the visualization of the similarity of acoustic features of the songs by applying a dimension reduction method. We selected UMAP as a dimension reduction method because the songs belonging to each category were most separated in the display space when we applied UMAP. The songs that the user listened to at home are concentrated at the left side of the visualization result as shown in Figure 5. On the other hand, the songs listened to at the university campus are concentrated on the right side. The songs that the user has not yet listened to are placed outside of them. This suggests the correlation between the location where the user listened to the songs and the acoustic features.

C. Visualization results and discussion

We observed the differences between music listened to at home and music listened to at the university campus using the interactive selection and emphasis functions of the visualization. Figures 8 and 9 show the visualization result limited to the songs listened to at home and another result limited to songs listened to at the university campus, respectively.

TABLE II
COLORMAP TABLE.

RED	Song user wants to listen to at home
BLUE	Song user wants to listen to at school
GRAY	Song user hasn't listened

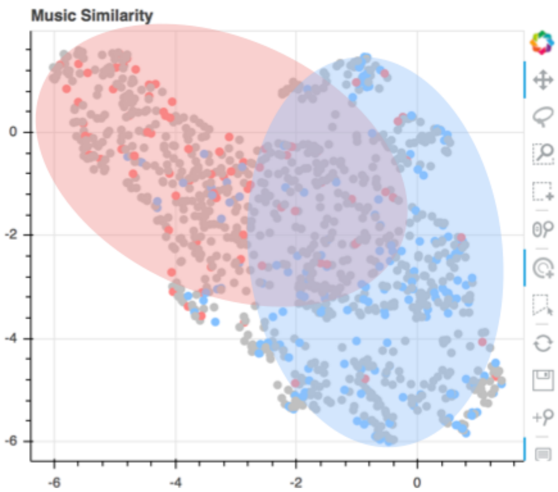

Fig. 5. Dimensionality reduction using Multi-Dimensional Scaling (UMAP).

Based on these results, we used Librosa [12] to check which acoustic feature values correlate to the songs listened to at each place. The most prominent features in the five acoustic feature values shown in Table III are "RMS energy" and "spectral contrast." Values of "RMS energy" increase from the upper left to the lower right, as shown in Figure 6. On the other hand, values of "spectral contrast" increase from the lower right to the upper left, as shown in Figure 7. We presume that many of the songs listened to at home tend to be songs with a small number of accompaniments (e.g. with just a guitar), because they have low "RMS energy" and high "spectral contrast". On the other hand, most of the songs listened to at the university campus have high "RMS energy" and low "spectral contrast", so we presume that we tend to be songs with many musical instruments.

TABLE III
LIST OF ACOUSTIC FEATURE VALUES UNDER ANALYSIS IN LIBROSA

tempo	Beats per minute
RMS energy	Root-mean-square (RMS) of power (Average volume)
spectral contrast	Strengths of spectral peaks
spectral flatness	Entropy of spetrum (Power of noise-like sounds)
spectral roll-off	Frequency that archives 85% of amount of spectrum

V. CONCLUSION AND FUTURE TASKS

This paper presented the visualization method and results for verifying the correlation between the locations where users

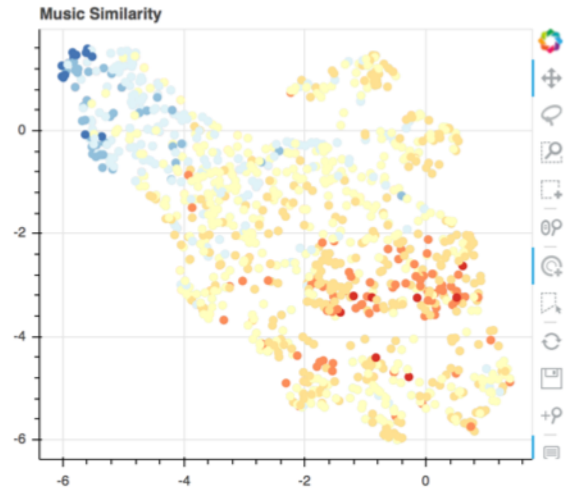

Fig. 6. Calculation result of "RMS energy" by Librosa.

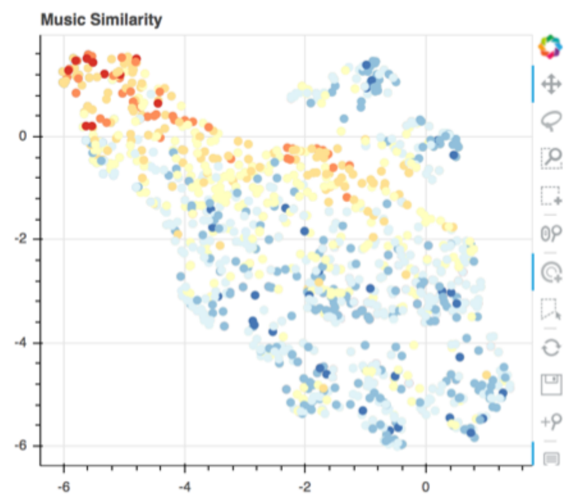

Fig. 7. Calculation result of "spectral contrast" by Librosa.

listened to the songs and their acoustic feature values. We found the correlation between the locations and the acoustic feature values from this visualization results.

We are discussing the following three issues as future work: visualization of the user's travel route, estimation of the moving speed, and selective song visualization in the scatterplot. Regarding the first and second issues, we would like to determine the destinations of the user's travel and the moving states of the travel. We would like to find the characteristics of the songs listened to at the destinations or each of the moving states. Here, we suppose four types of moving states based on the estimated speeds: walking, running, bicycle, and car/train. Regarding the third issue, we would like to solve a problem on a heavy overlap of the plots that may deteriorate the comprehensibility. We would like to develop an appropriate sampling of the songs to reduce the number of displayed plots and consequently reduce the overlap.

In addition, as a long-term problem, we would like to apply this correlation to classify songs that users prefer for each appreciation place by machine learning techniques. We aim to develop a music recommendation system that improves the satisfaction by applying the song classification technique. For this reason, recording more user data will be a very important issue in the future.

REFERENCES

[1] Shigeki Ohira and Katashi Nagao. Music player system adapted to walking/jogging rhythm. In *The 21st Annual Conference of the Japanese Society for Artificial Intelligence*, 2007.

[2] Hiromu Yakura, Tomoyasu Nakano, and Masataka Goto. Music recommendation system specialized for work bgm. In *IPSJ SIG technical reports, Vol. 112, No. 3*, 2016.

[3] Linas Baltrunas, Marius Kaminskas, Bernd Ludwig, and Omar Moling. Incarmusic: Context-aware music recommendations in a car. In *E-Commerce Web Technology, pp. 89–100*, 2011.

[4] Narumi Kuroko, Hayato Ohya, and Takayuki Itoh. A music recommendation system considering the location-based ordinariness. In *IPSJ SIG technical reports*, 2019.

[5] Adrian C. North, David J. Hargreaves, and Jon J. Hargreaves. Uses of music in everyday life. In *An Interdisciplinary Journal, Vol. 22, No. 1, pp. 1-77*, 2004.

[6] Anthony G. Greenwald, Timothy C. Brock, and Thomas M. Ostrom. Psychological foundations of attitude. In *Academic Press*, 1973.

[7] Gordon Reynolds, Dan Barry, Ted Burke, and Eugene Coyle. Interacting with large music collections: Towards the use of environmental metadata. In *IEEE International Conference on Multimedia and Expo, pp. 989-992*, 2008.

[8] Ian T. Jolliffe and Jorge Cadima. Principal component analysis: a review and recent developments. In *The Philosophical Transactions of the Royal Society*, 2016.

[9] Laurens van der Maaten and Geoffrey Hinton. Visualizing high-dimensional data using t-sne. In *journal of machine learning research, pp. 9:2579–2605*, 2008.

[10] Maa Cox and Tf Cox. Multidimensional scaling. In *Handbook of data visualization*, 2008.

[11] Leland Mcinnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. In *arXiv preprint arXiv:1802.03426*, 2018.

[12] Brian McFee, Colin Raffel, Dawen Liang, Daniel P.W. Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto. librosa: Audio and music signal analysis in python. In *Proc. of the 14th Python in Science Conf. (SCIPY), pp. 18-24*, 2015.

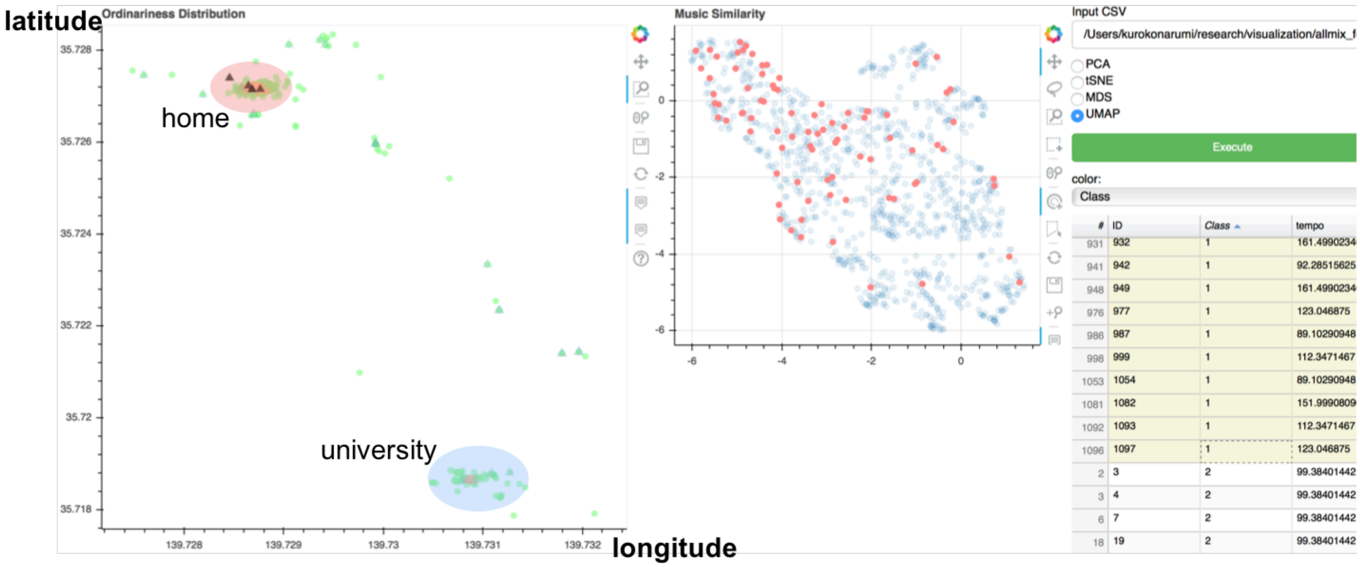

Fig. 8. Visualization limited to music listened to at home.

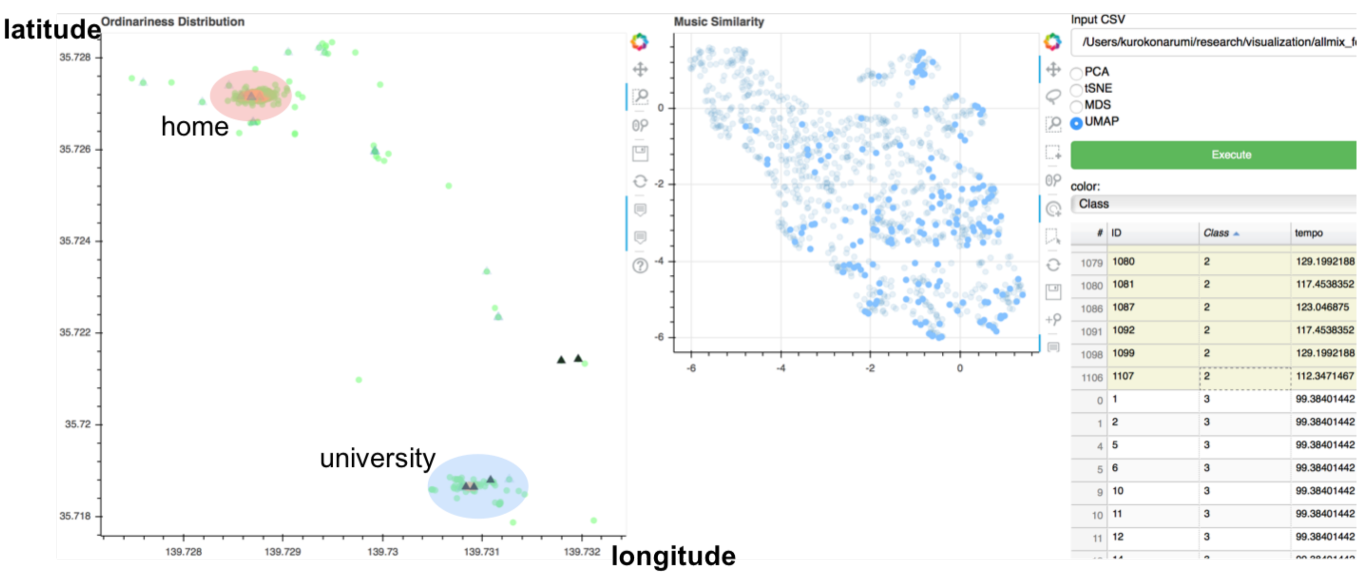

Fig. 9. Visualization limited to music listened to at university.