2009 International Symposium on Nonlinear Theory and its Applications
NOLTA'09, Sapporo, Japan, October 18-21, 2009

Music Similarity Measure using a Nonlinear Time Series Analysis Method

Ken'ichi Sawai†, Yoshito Hirata‡ and Kazuyuki Aihara‡†

†Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo
‡Institute of Industrial Science, University of Tokyo
Email: ken1@sat.t.u-tokyo.ac.jp

Abstract—This paper describes how to apply a nonlinear time series analysis method to musical scores. The technique is called cross recurrence plot. This can measure the similarity between two time series data. In our study, we use symbolic music information as obtained from musical scores. By regarding the symbolic information as time sequences, we can consider applying cross recurrence plots.

When we use cross recurrence plots, it is necessary to define the distance between two time points. If the data to be analyzed were numerical, the distances could be easily defined as Euclidean distances or infinity norms, for example. However, since musical data are the objects of the analysis, a new distance function is needed. Therefore, we propose a kind of edit distances by extending the distance for neuronal impulse trains. By using cross recurrence plots, we can quantify the atmospheres of musical pieces.

1. Introduction

There are some identical atmospheres of each musical songs even if they are played by different players. This suggests that there are some characteristics in the musical scores. Our goal of this study is to extract the features of scores. For this purpose, we propose a method for measuring music similarity between two scores.

Recent developments in computer technology have allowed us to have an large amount of music information in each own computer or through the Internet. Hence a lot of music researches are described in the context of music information retrieval. Melodic similarity is one of the main research fields of music information retrieval of both symbolic data (e.g. [3, 4]) and audio data (e.g. [6]). Many researches of similarity are proposed and almost all methods are measuring a rather small number of notes.

In addition to the above works, there are many other works related to ours, such as the application of dynamical systems theory to music [1]. Our methods can be also available for music information retrieval systems as well as other studies on similarity. And we think that our methods may be also useful for helping musicologists to analyze the relations among the songs.

2. Introduction of Recurrence Plots

2.1. Recurrence plots

Recurrence plot is a method for nonlinear time series analysis [2]. A recurrence plot is a 2-dimensional plot that visualizes recurrences in a time series $\{\boldsymbol{x}(i)\}$ for time $i = 1, 2, \ldots, N$, where N is the length of the time series. In a recurrence plot, both axes are time indexes. Given a distance function d and a pre-defined threshold ϵ, we plot a point at (i, j) if two states are close to each other. That is,

$$\boldsymbol{R}_{ij} = \Theta(\epsilon - d(\boldsymbol{x}(i), \boldsymbol{x}(j))), \quad i, j = 1, ..., N, \tag{1}$$

$$\Theta(u) = \begin{cases} 1 & (u > 0), \\ 0 & \text{otherwise}, \end{cases} \tag{2}$$

where $\boldsymbol{R}_{ij} = 1$ means plotting (i, j). The plotted points are called recurrence points. As a result, the points on the central diagonal line of a recurrence plot are all plotted and a recurrence plot is symmetric with respect to the central diagonal line.

2.2. Cross Recurrence Plots

Cross recurrence plot [8] is a bivariate extension of recurrence plots. We can obtain a cross recurrence plot between time series $\{\boldsymbol{x}(i)\}$ and $\{\boldsymbol{y}(j)\}$ by the next equation instead of Eq (1):

$$\boldsymbol{CR}_{ij} = \Theta(\epsilon - d(\boldsymbol{x}(i), \boldsymbol{y}(j))), \qquad i = 1, ..., M,\ j = 1, ..., N, \tag{3}$$

where $\boldsymbol{CR}_{ij} = 1$ means plotting (i, j). It is not necessary for M and N, the sizes of two time series data, to be the same. By observing a cross recurrence plot, we can examine the relation between two time series data.

2.3. Recurrence Quantification Analysis

There are some quantities calculated from recurrence plots. In this paper, we use the following two measures [7]:

- RR (recurrence rate)
  The most simplest measure is RR which is the density of recurrence points, i.e.
$$RR = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{R}_{ij}. \tag{4}$$
  RR represents the likelihood that a state recurs to its similar state.

- DET (determinism)
  DET is the ratio of the number of the recurrence points forming the diagonal lines to the number of all recurrence points. Formally, by using $P(l)$, the histogram of diagonal lines with length l, or
$$P(l) = \sum_{i=1}^{N}\sum_{j=1}^{N}(1-\boldsymbol{R}_{i-1,j-1})(1-\boldsymbol{R}_{i+l,j+l}) \times \prod_{k=0}^{l-1}\boldsymbol{R}_{i+k,j+k}, \tag{5}$$
  DET is calculated by
$$DET = \frac{\sum_{l=l_{min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)}, \tag{6}$$
  where we consider that a recurrence point forms a diagonal line if at least l_{min} consecutive recurrence points are arranged diagonally.

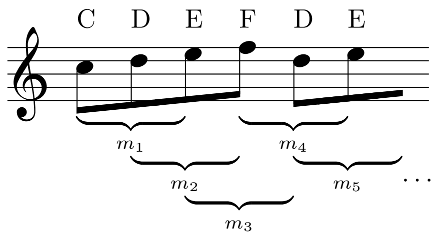
Quantifying the characteristics of recurrence plots as above is called *recurrence quantification analysis*. The above quantities can be also calculated for cross recurrence plots by replacing $\boldsymbol{R}$ with $\boldsymbol{CR}$ in each equations.

3. Music Similarity Measure using Cross Recurrence Plots

A musical score has variable information, for example, pitch, duration, tempo, and loudness. Here we adopt only pitches and onset times to consider music similarity because of their objectivity. The notes with the same pitch and duration imply variable meanings according to their contexts. Therefore, we regard a sequence of notes with a certain length of n as one state. From this, we can consider musical scores as time series data (see Fig. 1).

Recurrence plots need a distance function defined between two time points. Thus we discuss how to define the distance function.

The two scores with more similar atmospheres are thought to have similar melodies more. Hence, we consider of measuring the total similarity of two scores by using cross recurrence plots with a melodic similarity measure as a distance function.



m_i $(i = 1, 2, \dots)$: time series data from a score.

Figure 1: A musical score as time series data ($n = 3$).

There are many measures of melodic similarity proposed. The edit distance is one of the main types of the methods for measuring the melodic similarity [3]. The edit distance is defined as the minimum sum of the edit operation costs needed to transform one melody into the other, where an operation is an insertion, deletion, or substitution of a note.

We propose a kind of edit distances for measuring a distance between two sequences of notes by adjusting the distance between neuronal impulse trains [5]. While other methods of distances are also available, here we propose a new similarity which is measured with pitches and rhythms because we try to ignore the tempo variations. The method described in [5] is that,

1. set the cost for inserting and deleting a spike to 1, and set the cost for moving a spike a unit time to c_t. That is, when a temporal difference of two spikes is represented as Δt, the moving cost is equal to $c_t\Delta t$;
2. convert one of the two spike trains into the other by inserting, deleting or moving spikes with the costs;
3. find the conversion which achieves the minimum sum of the costs and regard the sum as the distance between the spike trains.

The brief image of conversion is shown in Fig. 2, where the horizontal axis shows time, the vertical solid lines represent the time of the spikes, the solid arrows mean insertion or deletion, and the dotted arrows correspond to moving actions. The minimum sum of the costs is computed by the dynamic programming.

This method can evaluate the difference of rhythm patterns of two melodies. To consider the pitch differences as well, we modify the above method by replacing the cost for moving a spike with the cost for moving a note. The cost for moving a note is defined as the weighted sum of the temporal differences and the pitch differences. That is, when the temporal difference and the pitch difference of two notes are Δt and Δp respectively, the moving cost is $c_t\Delta t + c_p\Delta p$, where c_t and c_p are the weights. Differences of pitches

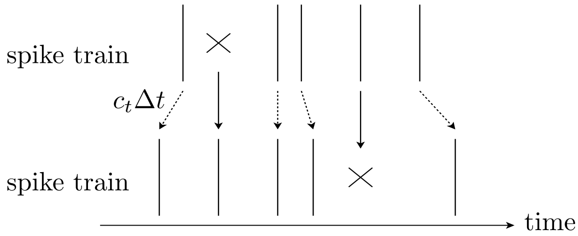


Figure 2: The brief image of the edit distance between spike trains [5].

are measured by semitones. This modification is explained in Fig. 3, where the horizontal axis is time and the vertical shows the pitch, the coordinates represent the notes, the solid arrow is deletion and insertion (there is no inserting action in this case), and the dotted arrows correspond to moving actions.

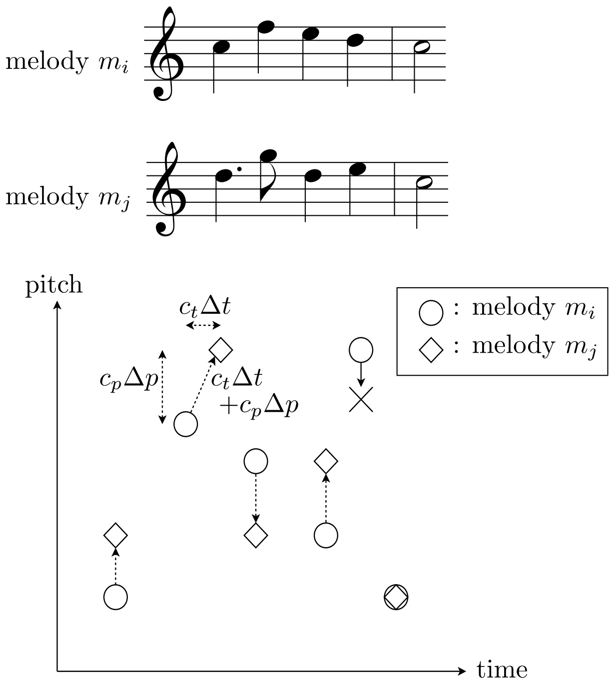


Figure 3: The brief image of the distance between melodies.

When calculating the distance, we translate one of two melodies so that the pitches of the initial notes are the same. This is because we perceive the melodies in different keys as the same. Moreover, in order to take into account the difference of time scales, we normalize a sequence of n consecutive notes so that the onset time of the first note is 0 and the onset time of the last note is 1.

These above transformations mean that we treat the melodies in Fig. 4 as the same, for example. By the transformations, we can measure the local similarity more finely.

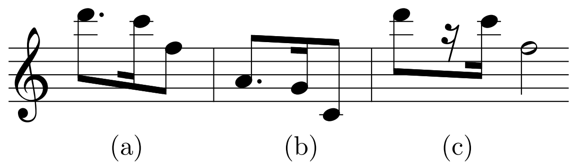


Figure 4: Examples of melodies treated as the same in our similarity method.

4. Results

We used the three scores of music:

Ernesto Köehler: Easy Exercise 1, 5 and 6 from "*35 Exercises for Flute Op. 33*",

and we call these scores as S_1, S_2 and S_3, respectively, after this.

We think that the scores S_1 and S_2 sound similar, and the score S_3 is different from the others. In addition, S_1 and S_2 are in major key while S_3 is in minor key. To quantify these impressions, we apply cross recurrence plots to the scores.

The results on applying cross recurrence plots with the state length of 4 to the scores S_1, S_2 and S_3 are shown in Fig. 5–7. Besides this, the values of RR and DET of the cross recurrence plots at the same state length are shown in Table 1. It can be observed that the values of RR and DET of the pair of the similar musical pieces are highest of the three pairs can be observed.

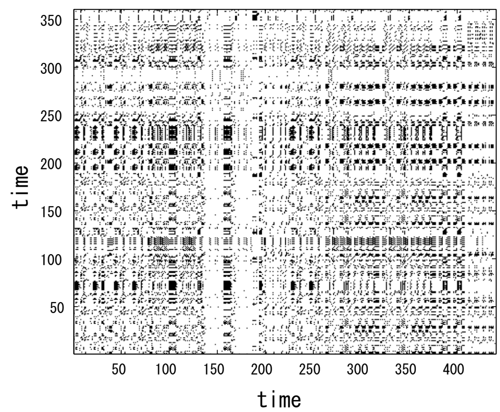


Figure 5: The melodic cross recurrence plot between S_1 and S_2.

5. Conclusion

A lot of music similarity measures have been proposed until now, but almost all of them were intended to be used for music information retrieval systems with queries by humming. Thus, the similarity measures used in such above systems are for short sequences of

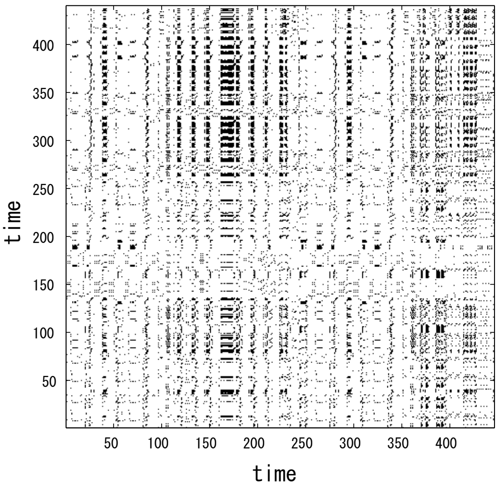


Figure 6: The melodic cross recurrence plot between S_2 and S_3.

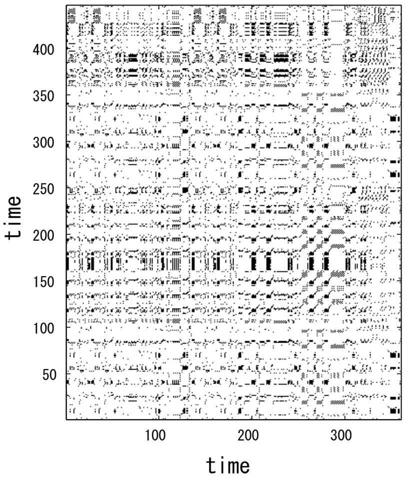


Figure 7: The melodic cross recurrence plot between S_3 and S_1.

Table 1: Feature quantities obtained from each cross recurrence plots.

combinations of scores	RR	DET
a-b	0.192	0.492
b-c	0.100	0.378
c-a	0.087	0.374

notes. We have suggested a possibility for our method to quantify the atmospheres of musical pieces.

In our method, other melodic similarity measures can be substituted for the distance function used in this method. The similarity measures used in music information retrieval systems often employ rough music information, such as the information of whether the pitches of notes are higher or lower than those of the previous notes. Because of threshold, it is not necessary to measure the melodic similarity in detail. Therefore, it may be also good to employ such a rough similarity method in cross recurrence plots.

Acknowledgment

This research is partially supported by Grant-in-Aid for Scientific Research (A) (20246026) from MEXT of Japan.

References

[1] Jean Pierre Boon and Olivier Decroly. Dynamical systems theory for music dynamics. *Chaos*, 5(3):501–508, 1995.

[2] Jean-Pierre Eckmann, Sylvie Oliffson Kamphorst, and David Ruelle. Recurrence plots of dynamical systems. *Europhysics Letters*, 4(9):973–977, 1987.

[3] Maarten Grachten, Josep-Lluís Arcos, and Ramon López de Mántaras. Melodic similarity: Looking for a good abstraction level. In *5th International Conference on Music Information Retrieval*, pages 210–215, 2004.

[4] Ning Hu, Roger B. Dannenberg, and Ann L. Lewis. A probabilistic model of melodic similarity. In *Proceedings of the International Computer Music Conference*, 2002.

[5] Jonathan D. Victor and Keith P. Purpura. Metric-space analysis of spike trains: Theory, algorithms and application. *Network: Computation in Neural Systems*, 8(2):127–164, 1997.

[6] Erling Wold, Thom Blum, Douglas Keislar, and James Wheaton. Content-based classification, search, and retrieval of audio. *IEEE Multimedia*, 3(3):27–36, 1996.

[7] Joseph P. Zbilut and Jr. Charles L. Webber. Embeddings and delays as derived from quantification of recurrence plots. *Physics Letters A*, 171(3-4):199–203, 1992.

[8] Joseph P. Zbilut, Alessandro Giuliani, and Jr. Charles L. Webber. Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification. *Physics Letters A*, 246(1-2):122–128, 1998.