1-33

Nonlinear Time Series Analysis Using Attractor Dimensionality Reduction with Neural Network

Ryosuke Shimizu Furugaki Tomiyuki Yoko Uwate Yoshifumi Nishio
(Tokushima University)

1. Introduction

It is difficult to analyze complicated time series data using Neural Network (NN). Therefore, it is important to search for the effective features of the data. In this study, we reduce the dimensionality of the data and search for features that are suitable for NN classification. We investigate the effect of dimensionality reduction methods on accuracy.

2. Proposed Method

In this study, the method of feature transformation for using 3-dimensional features of time series data is presented. (Fig. 1).

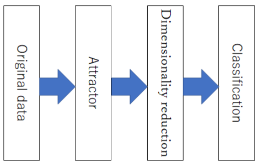

Figure 1: Feature translation.

Attractors are constructed by converting 1-dimensional data (1D-data) into 3-dimensional data using time delay embedding. $x(n)$ is the amplitude at a certain time. Equation (1) represents this system with the time delay value τ.

$$f(x) = [\ x(n)\ x(n+\tau)\ x(n+2\tau)\ ...] \quad (1)$$

In this study, the dimensionality reduction methods which are (a) Locally Linear Embedding (LLE) , (b) Principle Component Analysis (PCA) and (c) t-SNE are used. LLE is a method that uses the geodesic distance of multidimensional data. PCA is a method of multivariate analysis that synthesizes variables called principal components. t-SNE expresses proximity as a probability distribution. PCA is used for dimensionality reduction on linear data. LLE and t-SNE are used for dimensionality reduction on nonlinear data. The procedures of the methods are shown in the following steps.

(a) LLE

step 1. Determine the neighbors of each point.

step 2. Compute set weights that can reconstruct a point.

step 3. Compute lower-dimensional embedding.

(b) PCA

step 1. Compute a correlation matrix from a dataset.

step 2. Get the subspace as eigenvectors by singular value decomposition.

step 3. Compute lower-dimensional embedding.

(c) t-SNE

step 1. Determine the neighbors of each point.

step 2. Use t-distribution for distance distribution.

step 3. Compute lower-dimensional embedding.

3. Simulation Result

In this study, electromyogram (EMG) time series data are used as input. The experiments for collecting data consisted of freely and repeatedly grasping of different items which were essential to conduct the hand movements. In this study, three motion data are used. Data 1 is Spherical for holding spherical tools. Data 2 is Tip for holding small tools. Data 3 is Palmar for grasping with palm.

Table 1: The number of the time series data.

data	data1	data2	data3
train data	75	75	75
test data	75	75	75

Table 2 and 3 show the test accuracy of the conventional and the proposed methods using 1D-CNN. By setting the learning rate to 0.001 and the epoch of the training to 20, the training accuracy has reached a maximum value of 1.000. Test accuracy is calculated by averaging of 10 times. The test accuracy of applying PCA is the best, and that of applying t-SNE is the worst. LLE is also highly accurate. It can be said that the features of the data have changed significantly when converting the original data into the data using PCA.

Table 2: Test accuracy of the conventional method.

	test accuracy(%)
original	64.35

Table 3: Test accuracy of the proposed method.

	test accuracy(%)
LLE	72.00
PCA	74.62
t-SNE	52.22

4. Conclusion

In this study, the influence of the input features into accuracy is observed each dimensionality reduction. By transforming the data into attractor, it was found that the complex nonlinear data had a certain correlation. It turns out that t-SNE lost the data connection and erased the feature. In the future, it is necessary to investigate the reason for the large change in accuracy depending on the choice of dimensionality reduction.