Recurrent Self-Organizing Map implemented to detection of temporal line-movement patterns of *Lumbriculus variegatus* (Oligochaeta: Lumbriculidae) in response to the treatments of heavy metal

K.-H. Son¹, C. W. Ji², Y.-M. Park¹, Y. Cui³, H. Z. Wang³, T.-S. Chon² & E. Y. Cha¹
¹Department of Computer Science and Engineering, Pusan National University, Republic of Korea
²Division of Biological Sciences, Pusan National University, Republic of Korea
³Institute of Hydrobiology, Chinese Academy of Sciences, People’s Republic of China

Abstract

Measurement of behavioral responses have been recently considered as an important method for monitoring risk assessment. Computational processing could be applied to continuous data for automatic determination of changes in behavioural states of indicator specimens. Behavioral monitoring could be used as an alternative tool to fill the gaps between large (e.g., ecological survey) and small (e.g., molecular analysis) scale methods for risk assessment. While the points were conventionally used for indicating movement of test specimens, the line shapes of blackworms, *Lumbriculus variegatus*, were trained by Artificial Neural Networks in this study. We proposed an unsupervised temporal model, Recurrent Self-Organizing Map (RSOM), to detect sequential changes in the line-movement of blackworms after the treatments of a toxic substance, copper, in this study. RSOM was feasible in addressing the stressful behaviors of indicator specimens such as body contraction, high degree of folding, etc. We demonstrated that the unsupervised temporal model is efficient in classifying temporal behavior patterns and could be used as an alternative tool for the real-time monitoring of toxic substances in aquatic ecosystems in the future. Keywords: Recurrent Self-Organizing Map, response behavior, temporal sequence processing, *Lumbriculus variegatus*. 
1 Introduction

Considering the urgency of water contamination, development of methods for assessing toxicity has been regarded as an important issue in maintaining sustainable ecosystem health. Toxicity exposed to ecosystems can be assessed in various scales. Previous practices in assessment, however, were skewed to either extremely large (e.g., biodiversity evaluation in communities) or small (e.g. molecular or chemical analyses) scales.

Recently, continuous behavioral monitoring of indicator organisms [1–4] has been considered as one of the efficient tools filling the gaps between the large and small-scale assessment methods. Monitoring of the locomotory behavior has been introduced as an efficient means of evaluation of contaminated ecosystems with toxic chemicals [5]. A numerous accounts of research on effects of toxic chemicals on behaviors of organisms have been reported in various taxa, including crustaceans [6, 7], snails [8], fish [9, 10] and insects [11, 12]. However, these studies have been mostly limited to observation of single or combinations of single behaviors mainly with qualitative descriptions. Quantitative characterization of behaviors, however, is difficult for analysis due to complexity residing in the behavioral data.

Theoretical study on behavior has been initiated with research on biological motion regarding random walk [13], correlation function [14, 15], move length analysis [16], fractal dimension [17–19], etc. However, these parameters are highly condensed and tend to emphasize the totality of the movement states of indicator organisms. Addressing behavioral states in a compressed form (i.e., as a parameter), however, may not be suitable for uniquely characterizing various behavioral patterns. Local information on movement patterns may also be critical in determining various states of animal behaviors [1, 2].

In this study we used Artificial Neural Networks (ANNs) to extract information from complex behavioral data. ANNs has been regarded as an efficient non-linear filter, and have been widely used for forecasting and data organization in ecological sciences [20–22]. Recently ANNs have been applied to behavioral monitoring. Kwak et al [1] implemented Multi-Layer Perceptron (MLP) for detecting response behaviors of medakas treated with Diazinon. Wavelets and ANNs have been also used in combination to detect changes in response behaviors of chironomids for water quality monitoring [23]. In addition to supervised learning by MLP, Self-Organizing Map (SOM) was implemented to patterning the movement tracks of indicator organisms in an unsupervised manner in response to the treatments of toxic substances. SOM was efficient in classifying different states of response behaviors of indicator organisms such as cockroach [3] and medaka [24].

In the previous studies with ANNs, however, the static patterns were mainly considered as input, and location of the specimens (i.e., points) was the main source for training the movement tracks in the networks. In this study we demonstrated the feasibility of SOM in revealing the patterns of sequential movement of indicator specimens in a recurrent manner. In addition, a species in
oligochaetes was selected as indicator organisms, and the line shapes of the specimens were used as input data for training.

Recent models of neural networks have been used for temporal sequence processing (TSP). The temporal network was reported to be more feasible in learning time series data than conventional methods based on linear (e.g., AR and ARMA) and non-linear (e.g., NARMAX and MARS) statistical analyses [25]. In ecology, the Elman [26] and recurrent [27] networks have been used for predicting the time-series data for community dynamics. These networks, however, were mainly used for training with the templates (i.e., supervised learning). In real situations, however, there are numerous patterns in behavioural data. Consequently it is difficult to have all the pre-determined patterns be ready for training, considering that a huge amount of behavioural data could be accumulated through the real-time, continuous recording. Data mining would be desired to provide the overall scope of the behavioral data. In this regard, we incorporated the unsupervised network to accommodate the sequential line-movement data of specimens collected from continuous recording.

Temporal Kohonen Map (TKM) [28], being derived from the Kohonen Self-Organizing Map [29, 30], has been regarded as an efficient learning tool for TSP. In the TKM the involvement of the earlier input vectors in each unit is represented by using recursive difference. An unsupervised temporal model, Recurrent Self-Organizing Map (RSOM), was further proposed to provide more flexibly in dealing with the sequential data. RSOM was originally designed by Varsta et al [31] and can be presented as an enhancement of the TKM algorithm. While TKM does not directly use the temporal contextual information of input sequences in weight updating [31], direct learning of the temporal context is possible with RSOM. It allows model building using a large amount of data with only a little a priori knowledge. RSOM provided promising results in dealing with classification of temporal data with simple property [31–33].

In this study, we proposed RSOM for detecting temporal response behaviors of *Lumbriculus variegatus* treated with copper. We demonstrated that RSOM was feasible in patterning the sequential line-movement of oligochaetes after the treatments of toxic substances and efficiently characterize the stressful behaviors of the specimens.

2 Materials and methods

2.1 Observation system

The body shape of the test specimens of *Lumbriculus variegatus* was recorded by using an observation system consisting of an observation aquarium, a camera and software for an image recognition system. During the observation period, groups and individuals of *Lumbriculus* were placed in a glass aquarium (diameter: 9 cm), and their position was scanned from top view in 0.25 sec intervals using a CCTV camera (Kukjae Electronics Co. Ltd.; IVC-841®) for two days (one day before the treatments and one day after the treatments). The analog data captured by the camera were digitized by using a video overlay board.
(Sigmacom Co., LTD.; Sigma TVII®), and were sent to the image recognition system to digitize the line-movement of tested specimens. The software for detecting the specimens and other supporting mathematical programs were developed by the Neural Network and Real World Application Laboratory, Department of Computer Science and Engineering, Pusan National University. The stock populations were maintained in a glass tank, and were reared with artificial dry diet (Tetramin®) under the light condition (back light; twenty five 0.2W green diodes vertically located underneath the observation cage 10 cm apart) of 24 hours.

2.2 Computational method

Self-Organizing Map (SOM) is a vector quantization method to map patterns from an input space $V_i$ onto lower dimensional space $V_M$ of the map such that the topological relationships between the inputs are preserved [29, 30] to find the best matching unit $b$ in time step $t$ in the following equation:

$$\|x(t) - w_b(t)\| = \min_i \{\|x(t) - w_i(t)\|\}$$  \hspace{1cm} (1)

where $i \in V_M$, $x(t)$ is an input vector, and $w_i(t)$ is a weight vector of the unit $i$ in the map. Subsequently the weight vector of the best matching unit $b$ is updated towards the given input vector $x(t)$ according to

$$w_b(t+1) = w_b(t) + \gamma(t) h_b(t)(x(t) - w_b(t))$$  \hspace{1cm} (2)

where $\gamma(t), 0 < \gamma(t) \leq 1$, is a learning rate, and $h_b(t)$ is the neighbourhood function.

RSOM [31] is similar to the SOM except for the following difference equation, fig. 1:

$$y_i(t) = (1 - \alpha)y_i(t-1) + \alpha(x(t) - w_i(t))$$  \hspace{1cm} (3)

where $0 < \alpha \leq 1$ is a leaking coefficient, $y_i(t)$ is a leaked difference vector, $w_i(t)$ is the reference or weight vector in the unit $i$, and $x(t)$ is the input pattern in time step $t$. The best matching unit $b$ at time step $t$ is searched by

$$y_b = \min_i \{\|y_i(t)\|\}$$  \hspace{1cm} (4)

where $i \in V_M$. The process of updating weight is the same to SOM. However, the input sequence should be noticed before learning in a recurrent manner, fig. 1. The property of RSOM is described in reference articles in detail [31–33].
For input data, we obtained 13 $x$-$y$ coordinates of the line shape of blackworm specimens through computer recognition system measured in every 0.25 s interval. The coordinates were converted to a line consisting of 12 sub-segments with 12 lengths and 11 angles, fig. 2. These features carried information on the line shape of *Lumbriculus*. The general difference of the body shape was observable before and after the treatments, fig. 3. The lengths of the body segments were similar, being consistently shortened after the treatments, fig. 3(a). In contrast, the angles were smaller close to the center of the body, fig. 3(b). The angles in all sub-segments also decreased consistently after the treatments. This indicated that the body of the treated blackworms tended to contract and fold strongly after the treatments. For training RSOM, the data for 12 lengths and 11 angles were provided as input. The whole sequence of the line movement in two days (one day before the treatments and one day after the treatments) was divided into 100 sections with equal intervals (duration of ca 28.8 minutes). The line-data (12 lengths and 11 angles for the body segments, fig. 2) of the specimens at the beginning point was selected for the initial data for each section. Subsequently eleven more line data were selected in every 25 s interval in each section merged to the initial line-data. In total the 12 sequential line-data for 5 minutes (25 s X 12 = 300 s) were regarded as a sample unit provided to RSOM as input. Twenty specimens were observed for recording. Due to difficulty of detection of the line-shapes, however, some portion of data segments were not properly recorded due to noise. In this study we selected 6 specimens with the full records of the sequence movement during the whole observation period. The overall movement patterns in the specimens not selected for training were in general similar to the patterns of the specimens used for training according to the preliminary studies.

Figure 2: The body (bold line) of *Lumbriculus variegatus* consisting of 12 sub-segments with the lengths (e.g., $d_1$, $d_2$, etc) and the angles (e.g., $a_1$, $a_2$, etc)
3 Results

For the purpose of patterning the sequential line-movements of *Lumbriculus*, we initially checked angle and length separately for training RSOM. Subsequently length and angle were combined, and the whole data were trained with the network.

3.1 Movement patterns based on angles

The angle changes in the line-movements of the test specimens were accordingly grouped on RSOM, fig. 4(a), (b). According to the Ward’s linkage method calculated under MATLAB environment [34], the patterns were largely grouped into 4 patterns, fig. 4(a), (c). The horizontal gradient of the angle changes was observed from right to left area of the map. Angle changes in the body segments were minimal in the right area, while the specimens’ body shape was more folded in the left area as shown in, fig. 4(b).

The samples were divided to two main clusters on the map. In the right area, cluster B occupied a large area. In this cluster, the stretched line shape was mainly observed in the movement of the specimens, fig. 4(b). The segments before the treatments (white circle) were dominantly grouped in cluster B. The other cluster A occupied a broad area of the map at the left hand side. The folded body shapes were abundantly observed in this cluster. The cluster A was sub-clustered to smaller groups, AI and AII. The gradients were also observed between the sub-clusters. The highly folded ones appeared in the sub-cluster AI at the bottom left area of the map, while the data segments with less folded shapes were grouped in subcluster AII, fig. 4(b). The cluster AII was further divided to sub-sub-clusters, AIIa and AIIb. The segments in cluster AIIa in the
top area were presented by partial folding in the body segments, while the body segments in AIIb were characterized by the “U” shape, fig. 4(b). In cluster AI, the segments after the treatments (dark circle) were dominantly grouped. The segments in cluster AII were mixed between ‘before’ and ‘after’ the treatments.

Figure 4: (a) Grouping of the sequential line-movements of *Lumbriculus* specimens after training with RSOM based on angles of body segments (Clustering carried out on the patterned nodes by the Ward linkage method. White and black circles indicate the segments obtained from ‘before’ and ‘after’ the treatments respectively. Size of the circles indicates the number of line segments grouped in the RSOM units relatively (Max. number of the samples grouped in one unit; 120)). (b) Time sequence of the line-movements of *Lumbriculus* in different clusters listed on fig. 4(a) (c) Dendrogram of the RSOM units on fig. 4(a), by the Ward’s linkage method.
Figure 5: (a) Grouping of the sequential line-movements of *Lumbriculus* specimens after training with RSOM based on lengths of body segments (Clustering carried out on the patterned nodes by the Ward linkage method. White and black circles are explained in fig. 4 (Max. number of the samples grouped in one unit; 120)). (b) Time sequence of the line-movements of *Lumbriculus* in different clusters listed on fig. 5(a) (c) Dendrogram of the RSOM units in fig. 5(a) by the Ward’s linkage method.

### 3.2 Movement patterns based on lengths

The changes in length of the body segments of the blackworm specimens were also accordingly grouped on RSOM, fig. 5(a), (b). According to Ward’s linkage method, the patterns were grouped to 4 patterns, fig. 5(a), (c). In this case the gradient of the length was observed diagonally. From bottom left to top right on the map, the length of body segments accordingly decreased, fig. 5(a), (b). The patterned samples were divided to two large clusters, C and D. In cluster C occupying a large area in the bottom left area, the samples with long body segments were characteristically observed. Cluster C was further divided to two
sub-clusters, CI and CII, depending upon the degree of length. In cluster CI at the bottom left corner, the segment length was in the maximum range while the specimen’s body was close to the line-shape without much folding, fig. 5(a), (b). The body length in cluster CII was relatively more contracted than in cluster CI, fig. 5(b). In cluster CI, the body segments before the treatments were more abundantly observed. However, some body segments after the treatments also occurred in this cluster. In cluster CII, the body segments before the treatments were abundant but were mixed with the body segments after the treatment in some degree.

In cluster D close to the upper right area of the map, the contracted forms of body segments were strongly grouped. Especially in cluster DI, the specimens highly contracted and folded. In cluster DII, the body segments were less contracted. In cluster DI, the segments after the treatments were strongly grouped. The segments after the treatments were still dominant in cluster DII. It appeared that the degree of grouping was stronger for the body segments after the treatments (i.e., DI and DII) than before the treatments (i.e., CI and CII).

3.3 Movement patterns based on combination of angles and lengths

The data for angles and lengths of the body segments of the blackworm specimens were combined and were subsequently used for training RSOM, fig. 6(a), (b), (c). Clustering appeared in a characteristic manner more diversely compared with separate patterning by angles, fig. 4, and lengths, fig. 5. According to the Ward’s linkage method, the samples were divided to 6 groups, with inclusion of sub- and sub-sub-clusters, fig. 6(a), (c).

Although clustering was diverse, the gradient was still observed diagonally from bottom left to top right on the map. While the body segments with larger and less folded bodies were observed in cluster F in the area of bottom right, the shorter and strongly folded bodies were placed in cluster E in the area of top left, fig. 6(a), (b).

Sub-clusters were similarly divided according to degree of contraction and folding. In cluster FI at the bottom right corner, the longer and less curved specimens were observed. The sub-cluster FI was further divided according to the line shape of the specimens. While more straight forms were allowed in the body segments in sub-sub-cluster FIa, folding (e.g., the fifth snap shot in FIb, fig. 6b) was produced in the sequence of the line-movement in the sub-sub-cluster FIb.

Sub-cluster, FII, was differentiated from FI regarding the stronger degree of contraction and folding. The gradient was further observed in division of sub-clusters. In sub-sub-cluster FIa, the body segments were strongly contracted and folded. In sub-sub-cluster FIIb, however, the shape of the body segments was relatively longer and somewhat similar to the body segments shown in sub-cluster FI. Clear difference of the body segments between these two clusters needs more verification, but the body segments in cluster FIIb appeared to include the more contracted segments (e.g., the sixth snap shot in FIIb, fig. 6(b) during the course of sequential movement.
Figure 6: (a) Grouping of the sequential line-movements of *Lumbriculus* specimens after training with RSOM based on angles and lengths of body segments (Clustering carried out on the patterned nodes by the Ward linkage method. White and black circles are explained in fig. 4 (Max. number of the samples grouped in one unit; 100)). (b) Time sequence of the line-movements of *Lumbriculus* in different clusters listed on fig. 6(a) (c) Dendrogram of the RSOM units in fig. 6(a) by the Ward’s linkage method.

In cluster E, the samples showed the strongly contracted and folded bodies, fig. 6(b). In the body segments belonging to sub-cluster EI, the length was in the least range while the degree of folding was in the highest range. Almost all the grouped samples in this cluster were the line-movements after the treatments. In sub-cluster EII, the body was less contracted and the degree of folding decreased, but the samples were still dominated by the line-movement after the treatments. The sequential movement in the contracted and folded forms of the treated
specimens was accordingly revealed in clusters EI and EII. This demonstrated that RSOM would efficiently visualize the symptomatic movements of blackworms after exposure to copper. In contrast to cluster E, the samples before the treatments were more abundantly observed in the areas of cluster F (e.g., upper area in FIIb), however the segments after the treatments were also included in some degree. Overall, classification was more discernible with the body segments after the treatments for the combined data of angles and lengths.

4 Discussion and conclusions

In this study we used a species in oligochates, *Limnodurlus variegatus*, for continuous monitoring of temporal line-movement in response behaviors to the treatments of toxic substances. RSOM was feasible in classifying the sequential line-movements of the blackworms. To the best of our knowledge, no computational method has been carried out on patterning behaviors of blackworms with temporal movement, fig. 4, 5 and 6. This type of the line patterning would be useful for monitoring behavioral changes in animals with long body shapes (e.g., other annelids, snakes, eels, etc) in the future. If the lines are detected instead of the points, the scope in movement changes would be broader in providing diverse information on movement.

We analysed the response behaviors based on separate criteria, angle and length. Firstly, patterns of body folding were accordingly revealed by training with the angle data, fig. 4(b), by the trained RSOM. With the data for the lengths subsequently, fig. 5(b), the ‘line’ part was more clearly identified: the straight lines (e.g., CI in fig. 5(b)) were efficiently grouped. In the data for angles, fig 4(b), however, the straight segments shown in CI were not observed. Conversely, the highly folded ones (e.g., A1 in fig. 4(b)) were not observed in fig. 5(b). RSOM was efficient in extracting information from complex data accordingly to characteristics of input variables.

We used blackworms for indicator species for behaviour monitoring. Aquatic oligochaetes, such as *Lumbriculus*, are important taxa in freshwater aquatic communities. The oligochaetes serve diverse roles as aiding in decomposition of organic materials in the sediment. *Lumbriculus variegatus* has been proposed as a standard organism for sediment bioaccumulation tests [35]. Researchers reported that *Lumbriculus* have several locomotor behaviors such as rapid withdrawal, crawling, body reversal, and helical swimming. There are stereotyped behaviors that can be used for the sub-lethal toxicology [36, 37]. In this study RSOM was feasible in identifying some symptomatic sequence of line-movements of blackworms, fig. 4, 5 and 6, and the results from our study confirmed the toxicological effect of black worms on behaviours.

In this study we used copper as a toxic agent. Copper has commonly been used for fertilizers, and especially pesticides due to its antifungal properties [38]. Excessive levels can contaminate aquatic ecosystems [38, 39] even though it is an essential trace element. Sub-lethal copper exposure alters a number of behaviors in invertebrates. Locomotor behavior can also be adversely affected by copper exposure [40–43]. Bruce et al [44] reported that copper exposure can
significantly affect the ability of aquatic oligochaetes to avoid predators. We were able to demonstrate that copper would cause typical symptomatic behaviours in temporal sequence of line-movement of blackworms by using RSOM.

In conclusion, RSOM was feasible in addressing the sequential line-movements of stressful behaviors of blackworms such as body contraction, high degree of folding, etc. The temporal patterns of body shapes after the treatments were more strongly grouped than the temporal patterns before the treatments. The recurrent unsupervised model in ANNs used in this study would be an efficient tool for identifying temporal behavior patterns of indicator specimens and could be used as an in-situ real-time monitoring device in aquatic ecosystems in the future.

Acknowledgement

This subject is supported by Korea Ministry of Environment as “The Eotechnopia 21 project”.

Reference


[7] Roast, S.D., Widdows, J., Jones & M.B., Disruption of swimming in the hyperbenthic mysid Neomysis integer (Peracarida: Mysidacea) by the


