Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 22 (2013) 883 – 892

Procedia
Computer Science

17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems - KES2013

Extraction of Daily Life Log Measured by Smart Phone Sensors using Neural Computing

János Botzheim[a,b,*], Dalai Tang[a], Bakhtiar Yusuf[a], Takenori Obo[a], Naoyuki Kubota[a], Toru Yamaguchi[a]

[a]*Graduate School of System Design, Tokyo Metropolitan University, 6-6 Asahigaoka, Hino, Tokyo, 191-0065 Japan*
[b]*Department of Automation, Széchenyi István University, Egyetem tér 1, Győr, 9026, Hungary*

Abstract

This paper deals with the information extraction of daily life log measured by smart phone sensors. Two types of neural computing are applied for estimating the human activities based on the time series of the measured data. Acceleration, angular velocity, and movement distance are measured by the smart phone sensors and stored as the entries of the daily life log together with the activity information and timestamp. First, growing neural gas performs clustering on the data. Then, spiking neural network is applied to estimate the activity. Experiments are performed for verifying the effectiveness of the proposed method.

Keywords: smart phone sensors; daily life log; information extraction; neural computing; growing neural gas; spiking neural networks

1. Introduction

The emerging synthesis of information technology, network technology, and robot technology is one of the most promising approaches to realize a safe, secure, and comfortable society for the next generation [1]. An important thing in the technologies and methods is the structuralization of the information. By structuralization one can give qualitative meaning to the data which is useful in improving the accessibility and usability of information. Huge datasets can be obtained by sensor networks however useful, meaningful and valuable information should be extracted from such data.

Smart phones and tablet PCs are becoming more and more popular nowadays and their price is decreasing year by year. They can be equipped with various sensors such as gyro, accelerometer, illumination sensor, compass, camera and other sensors. Elderly people are the main target group of our research, since in Japan the increase of elderly people has become an important problem. Moreover, in the aging society the number of elderly people living alone or separated from their children is increasing. Elderly people unfamiliar with information home appliances also have easily started using tablet PC, because touch panels or touch interface have been popularized at ticket machines and information services in public areas.

*Corresponding author. Tel.: +81-42-585-8441 ; fax: +81-42-585-8441.
E-mail address: botzheim@sd.tmu.ac.jp.

doi:10.1016/j.procs.2013.09.171

Since smart phone is almost always with us it is easy to collect data by it. Daily life log can be created which can contain the time series of the measured sensory data together with activities information entered by the user. When we understand the information from the time series of data then it is useful for supporting the user. For example in case of elderly people, a robot partner can warn the person if he or she does unusual activity, or if the order of his or her activity is not normal, or shows irregular pattern based on the previous days' data.

Intelligent methods such as neural computing can be used to memorize patterns, sequences, and other information can be found also in the life log. In our previous paper we used smart phone to estimate human transport modes by spiking neural networks and evolution strategy [2]. In that case the neural network was trained by evolutionary computation in supervised manner.

In this paper we apply growing neural gas (GNG) and spiking neural network (SNN) for memorizing the daily life log and estimating the activities. In this case we use unsupervised learning for the neural networks. We used the combination of GNG and SNN also for gesture recognition in [3].

The paper is organized as follows. Section 2 presents informationally structured space and sensory inputs from smart phone. In Section 3 the proposed approach is discussed. Experiments are shown in Section 4. Conclusions are drawn in Section 5.

2. Information Support by Daily Life Log

2.1. Informationally Structured Space

Information resources and the accessibility within an environment are essential for both people and robots. Therefore, the environment surrounding people and robots should have a structured platform for gathering, storing, transforming, and providing information. Such an environment is called informationally structured space [4, 5]. The intelligent technology for the design and usage of the informationally structured space should be discussed from various points of view such as information gathering of real environment and cyber space, structuralization, visualization and display of the gathered information. The structuralization of informationally structured space realizes the quick update and access of valuable and useful information for people. It is very useful for both robots and people to easily access the information on real environments [6]. The information is transformed into the useful form suitable to the features of robot partners and people. Furthermore, if the robot can share the environmental information with people, the communication with people might become very smooth and natural.

Figure 1 illustrates the concept of informationally structured space when gathering data in daily life. In Fig. 1 different levels of information support can be seen. Personal information is gathered by an iPhone. Indoor life log can be produced by sensor network [7]. The next level is when the human activity log considers also outdoor information. In Fig. 1 a rechargeable IC card is shown which stores electric money can be used for purchasing and for public transport as a ticket. By this card we can trace the shopping and traveling activities of the person.

The collected information can be used also on different levels. The first level is the information support for family, or in case of elderly people for the caregivers. The other level is social network in the sense of a local community on the internet. The last level is the complete information support using internet.

In this research only smart phone is applied to collect personal information, indoor life log, and outdoor activity log. By understanding the daily log we can support the people by recognizing unusual patterns, differences between the daily logs on different days etc. In case of elderly people a robot partner can warn the person in case of irregular or strange behavior.

2.2. Sensory Inputs from a Smart Phone

An iPhone is equipped with various sensors such as gyro, accelerometer, illumination sensor, touch interface, compass, two cameras, and microphone [8]. The acceleration of human movement is calculated from the data measured by the accelerometer as follows:

$$a(t) = \sqrt{a_x(t)^2 + a_y(t)^2 + a_z(t)^2}, \tag{1}$$

where $a_x(t)$, $a_y(t)$, and $a_z(t)$ are the components of the acceleration in the unit directions at time t.

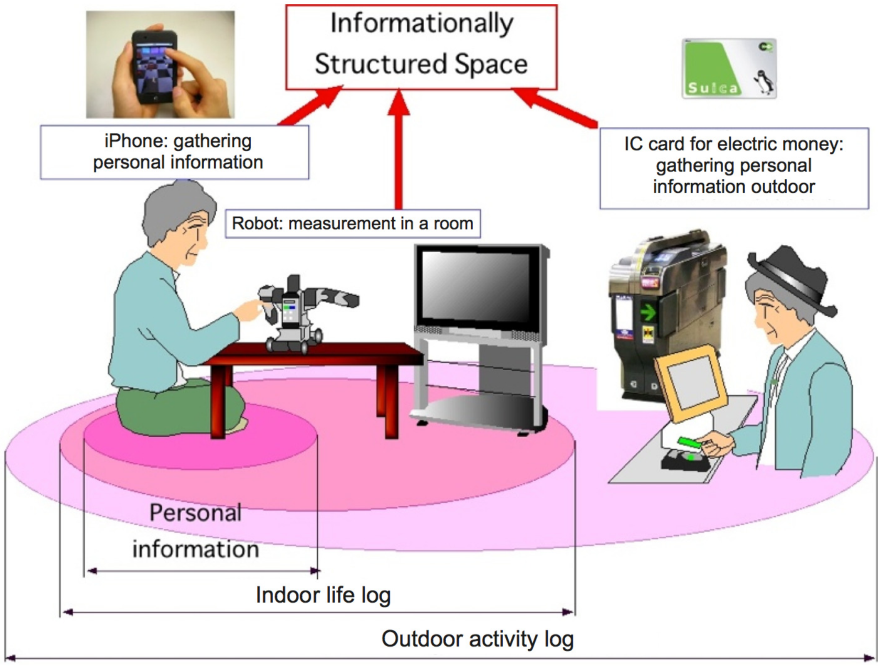

Fig. 1. Data gathering in daily life

The angular velocity is calculated from the data measured by the gyro sensor as follows:

$$v(t) = \sqrt{v_x(t)^2 + v_y(t)^2 + v_z(t)^2}, \tag{2}$$

where $v_x(t)$, $v_y(t)$, and $v_z(t)$ are the roll, pitch, and yaw angles at time t, respectively.

The movement distance is calculated by GPS as:

$$s(t) = \sqrt{(g_x(t) - g_x(t-1))^2 + (g_y(t) - g_y(t-1))^2}, \tag{3}$$

where $g_x(t)$, and $g_x(t-1)$ are the latitude components and $g_y(t)$, and $g_y(t-1)$ are the longitude components at time t and $t-1$, respectively. The altitude component of the GPS data is not considered here because of the different scale of that component.

After calculating the values of acceleration, angular velocity, and movement distance according to a given sampling time, their values are normalized. Figure 2 illustrates a sample data of the measured user motion outdoor. The horizontal axis refers to the time, and the vertical axis means the normalized values of the measured signals after preprocessing by Eqs. (1), (2), (3). The gray line is the data measured by the accelerometer. The blue line depicts the angular velocity calculated from the data measured by the gyro sensor. The green line is the moving distance calculated by the GPS. The GPS sensor has role mainly in outdoor situations which is illustrated in Fig. 2.

In this research the measured data by iPhone sensors are used for creating daily life logs. Several activities and behaviors of the person are considered. Anytime the person can record his or her activity by an application displayed in Fig. 3. The person can start and stop the recording, and save the log in a file. The activities can be indoor or outdoor. The GPS sensor carries less information in case of indoor activities, however it is important in outdoor. Figure 3 illustrates the 24 activities realized in the application. During the recording the sensory data, the selected activity, and the timestamp are saved in the given file according to the sampling time.

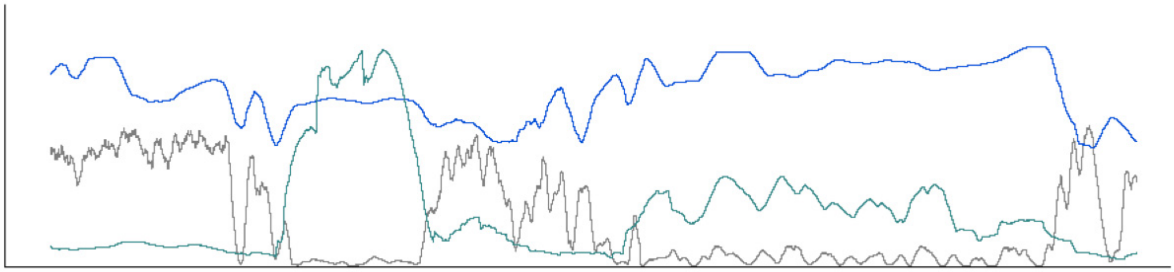

Fig. 2. Sample data in case of outdoor motion

(a) Illustration of activities on iPhone

outdoor	body care	meal	leisure	housework
stop	wake up	breakfast	TV	washing
walk	sleep	lunch	internet	cleaning
bicycle	bath	dinner	game	cooking
bus	wc	teatime	pray	shopping
train			hobby	chores
			leisure	

(b) List of activities

Fig. 3. Activities

3. Our Proposed Approach

The goal of the research is to extract the information from the daily life log. Neural computing is applied to learn the log. Our approach contains two steps. After preprocessing the sensory data by calculating the acceleration, angular velocity, and movement distance by Eqs. (1), (2), (3) at every timestamp, we use these data together with the activity information and apply clustering by growing neural gas in the first step. In the second step spiking neural network is applied to understand the spatiotemporal information and estimate the activity.

3.1. Growing Neural Gas for Information Extraction

Unsupervised learning is performed by using data without any teaching signals [9, 10, 11, 12, 13, 14, 15]. Self-organized map (SOM), neural gas (NG), growing cell structures (GCS), and growing neural gas (GNG) are well known as unsupervised learning methods. Basically, these methods use the competitive learning approach. The number of nodes and the topological structure of the network in SOM are designed beforehand [9, 10]. In NG, the number of nodes is fixed beforehand, but the topological structure is updated according to the distribution of sample data [11]. On the other hand, GCS and GNG can dynamically change the topological structure based on the adjacent relation (edge) referring to the ignition frequency of the adjacent node according to the error index. However, GCS does not delete nodes and edges, while GNG can delete nodes and edges based on the concept of ages [14, 15]. Furthermore, GCS must consist of k-dimensional simplices whereby k is a positive integer chosen in advance. The initial configuration of each network is a k-dimensional simplex, e.g., a line is used for $k = 1$, a triangle for $k = 2$, and a tetrahedron for $k = 3$ [12, 13]. GCS has been applied to construct 3D surface models by triangulation based on 2-dimensional simplex. However, because the GCS does not delete nodes and edges, the number of nodes and edges is over increasing. Furthermore, GCS cannot divide the sample data into several segments.

In the learning algorithm of GNG [14, 15] the following notations are used:

- w_i: the n dimensional vector of a node ($w_i \in \mathbb{R}^n$)
- A: set of nodes
- N_i: the set of nodes connected to the i-th node
- c: the set of edges
- $a_{i,j}$: the age of the edge between the i-th and the j-th node

The steps of the GNG algorithm are as follows:
Step 0. Generate two units at random position, w_{c1}, w_{c2} in $\mathbb{R}^n$. Initialize the connection set.
Step 1. Generate an input data v randomly according to $p(v)$ which is the probability density function of data v.
Step 2. Select the nearest unit (winner) s_1 and the second-nearest unit s_2 by:

$$s_1 = \arg\min_{i \in A} ||v - w_i|| \tag{4}$$

$$s_2 = \arg\min_{i \in A \setminus \{s_1\}} ||v - w_i|| \tag{5}$$

Step 3. If a connection between s_1 and s_2 does not exist already, create the connection. Set the age of the connection between s_1 and s_2 to zero:

$$a_{s_1,s_2} = 0 \tag{6}$$

Step 4. Add the squared distance between the input data and the winner to a local error variable:

$$E_{s_1} \leftarrow E_{s_1} + ||v - w_{s_1}||^2 \tag{7}$$

Step 5. Update the reference vectors of the winner and its direct topological neighbors by the learning rate η_1 and η_2, respectively, of the total distance to the input data:

$$w_{s_1} \leftarrow w_{s_1} + \eta_1 \cdot (v - w_{s_1}) \tag{8}$$

$$w_j \leftarrow w_j + \eta_2 \cdot (v - w_j) \quad \text{if } c_{s_1,j} = 1 \tag{9}$$

Step 6. Increment the age of all edges emanating from s_1:

$$a_{s_1,j} \leftarrow a_{s_1,j} + 1 \quad \text{if } c_{s_1,j} = 1 \tag{10}$$

Step 7. Remove edges with an age larger than a_{max}. If this results in units having no more emanating edges, remove those units as well.
Step 8. If the number of input signals generated so far is an integer multiple of a parameter λ, insert a new unit as follows:

i. Select the unit q with the maximum accumulated error.
ii. Add a new unit r to the network and interpolate its reference vector from q and its neighbor f with the largest error variable:

$$w_r = 0.5 \cdot (w_q + w_f) \tag{11}$$

iii. Insert edges connecting the new unit r with units q and f, and remove the original edge between q and f.
iv. Decrease the error variables of q and f by a fraction α:

$$E_q \leftarrow E_q - \alpha \cdot E_q \tag{12}$$
$$E_f \leftarrow E_f - \alpha \cdot E_f \tag{13}$$

v. Interpolate the error variable of r from q and f:

$$E_r = 0.1 \cdot (E_q + E_f) \tag{14}$$

Step 9. Decrease the error variables of all units:

$$E_i \leftarrow E_i - \beta \cdot E_i \quad (\forall i \in A) \tag{15}$$

Step 10. Continue with Step 1 if a stopping criterion (e.g. net size or some performance measure) is not yet fulfilled.

In our case the nodes (reference vectors) are five dimensional. They contain the three preprocessed sensory data, the activity information, and the timestamp. However, when the nodes are updated and they positions are changed according to Eqs. 8 and 9, only the three sensory data positions are changed, while the activity information and the timestamp remain unchanged.

3.2. Spiking Neural Network for Activity Estimation

Growing neural gas can learn the important topological relations in a set of data vectors. Several other types of artificial neural networks have been proposed to realize clustering, classification, non-linear mapping, and control [16, 17, 18, 19]. Basically, artificial neural networks are classified into pulse-coded neural networks and rate-coded neural networks from the viewpoint of abstraction level [18]. A pulse-coded neural network approximates the dynamics with the ignition phenomenon of a neuron, and the propagation mechanism of the pulse between neurons.

One important feature of pulse-coded neural networks is the capability of temporal coding. In fact, various types of spiking neural networks (SNNs) have been applied for memorizing spatial and temporal context [5, 16, 18]. We use a modified simple spike response model to reduce the computational cost.

The membrane potential, or internal state $h_i(t)$ of the i-th neuron at the discrete time t is given by:

$$h_i(t) = \tanh(h_i^{syn}(t) + h_i^{ref}(t) + h_i^{ext}(t)), \tag{16}$$

where $h_i^{syn}(t)$ includes the pulse outputs from the other neurons, $h_i^{ref}(t)$ is used for representing the refractoriness of the neuron, $h_i^{ext}(t)$ is the input to the i-th neuron from the environment. The hyperbolic tangent function is used to avoid the bursting of neuronal fires.

The first term $h_i^{syn}(t)$ is calculated as follows:

$$h_i^{syn}(t) = \gamma^{syn} \cdot h_i(t-1) + \sum_{j=1, j\neq i}^{N} w_{j,i} \cdot h_j^{PSP}(t-1), \tag{17}$$

where γ^{syn} is the temporal discount rate, $w_{j,i}$ is a weight from the j-th neuron to the i-th neuron, $h_j^{PSP}(t)$ is the presynaptic action potential (PSP) approximately transmitted from the j-th neuron at the discrete time t, and N is the number of neurons. When the internal state of the i-th neuron reaches the predefined threshold, a pulse is outputted as follows:

$$p_i(t) = \begin{cases} 1 & \text{if } h_i(t) \geq \theta, \\ 0 & \text{otherwise,} \end{cases} \tag{18}$$

where θ is a threshold for firing. When the neuron is fired, R is subtracted from $h_i^{ref}(t)$:

$$h_i^{ref}(t) = \begin{cases} \gamma^{ref} \cdot h_i^{ref}(t-1) - R & \text{if } p_i(t-1) = 1, \\ \gamma^{ref} \cdot h_i^{ref}(t-1) & \text{otherwise,} \end{cases} \tag{19}$$

where γ^{ref} is a discount rate and $R > 0$.

The presynaptic spike output is transmitted to the connected neuron according to PSP through the weight connection. The PSP is calculated as follows:

$$h_i^{PSP}(t) = \begin{cases} 1 & \text{if } p_i(t) = 1, \\ \gamma^{PSP} \cdot h_i^{PSP}(t-1) & \text{otherwise,} \end{cases} \tag{20}$$

where γ^{PSP} is a discount rate and $(0 < \gamma^{PSP} < 1)$. Therefore, the postsynaptic action potential is excitatory if the weight parameter, $w_{j,i}$ is positive. If the condition $h_j^{PSP}(t-1) < h_i^{PSP}(t)$ is satisfied, the weight parameter is trained based on the temporal Hebbian learning rule as follows:

$$w_{j,i} \leftarrow \tanh\left(\gamma^{wgt} \cdot w_{j,i} + \xi^{wgt} \cdot h_j^{PSP}(t-1) \cdot h_i^{PSP}(t)\right), \tag{21}$$

where γ^{wgt} is a discount rate and ξ^{wgt} is a learning rate.

We would like to estimate the person's activity at any time from the daily life log's sensory data and timestamp information using the reference vectors of GNG. In the SNN one neuron corresponds to one activity, thus the number of spiking neurons is equal to the number of activities. For each entry of the log using the entry's input information (the sensory data and the timestamp) the aim is to estimate the entry's output information (the activity) by the GNG and SNN. The external input of a given neuron in SNN will be calculated by the given entry's input information, in such a way, that we calculate the smallest Euclidean distance between the entry's input and those reference vectors (GNG nodes) which has the same output category (activity) as the given spiking neuron corresponds to. This smallest distance is transformed and used as the spiking neuron's external input. Smaller distance means higher external input. Formally:

$$h_i^{ext}(t) = \gamma^{ext} \cdot \left(1 - \min_k ||x(t) - w_i^{(k)}||\right), \tag{22}$$

where γ^{ext} is a scaling factor for the external input, $x(t) = (a(t), v(t), s(t), t)$ is the entry's input, and $w_i^{(k)}$ are those GNG nodes which belong to category i.

The output of the SNN is the estimation of the activity which is calculated by the winner neuron based on the PSP values. First, we check which neuron has the biggest PSP value for the given input data. If the winner PSP is larger than a predefined threshold, then the output of the SNN will be that activity which corresponds to this winner neuron. If the PSP is less than the threshold, then the output of SNN will be the same as in case of the previous input of the time sequence.

A common important point of the growing neural gas and spiking neural network is the Hebbian learning [20]. In GNG, Step 3 applies Hebbian learning since the correlation between the nodes is considered. In case of SNN the weight parameter is increased when simultaneous activation of neurons occur.

Fig. 4. The difference between acquired data (red) and computed data (black)

4. Experimental Results

In the experiment, we utilized the combination of GNG and SNN to produce more effective and efficient method for estimating the human daily life pattern. The experiment consists of two parts. First, we took the life log data using iPhone application developed by our team (Fig. 3). In addition to estimate human daily activity, through this experiment we want to investigate the variety of human daily life pattern according to cultural differences. For our purpose, we provide three different subjects from three different cultures, such as Indonesia, Mongolia and Japan.

Next, after acquired the required data from the subjects, we conduct data computation using GNG and SNN. In order to gain a good result we do some adjustments on the parameters. For this experiment, the parameters of GNG and SNN are shown in Tables 1 and 2, respectively.

Table 1. GNG parameters

Max. number of neurons	iterations	η_1	η_2	a_{max}	λ	α	β
300	40000	0.05	0.01	6	100	0.8	0.01

Table 2. SNN parameters

γ^{syn}	γ^{ref}	γ^{PSP}	γ^{wgt}	ξ^{wgt}	R	θ	output threshold	γ^{ext}
0.9	0.7	0.9	0.7	0.9	1	0.9	0.7	0.665

The result of the GNG and SNN computation is depicted in Fig. 4. According to Fig. 4 we can see that the acquired data as input data and the computed data produce only 0.15 average error, which proves that our method is sufficient enough to conduct the estimation. In Fig. 4 the horizontal axis refers to the time, while the vertical axis means the computed activity.

Different cultures produce different daily activities. In Fig. 5 and Fig. 6 the different characteristics of two subjects are displayed. Figure 5 shows that this subject has a breakfast daily activity, on the other hand, instead of breakfast subject 2 has lunch daily activity. Furthermore, as a religious person, subject 1 has a daily pray activity.

Compared to the previous day activity, subject 1 relatively has a similar daily activity pattern as illustrated in Fig. 7. Although he has similar daily activity, the result data show slight differences.

As a reference, we also provide the daily activity of a Japanese subject, especially the utilization of transportation (Fig. 8). As a Japanese living in big cities such as Tokyo, the train has been a favorite transportation.

In the experiments the number of data in the daily life log was about between 15000 and 30000. In Figs. 4-8 the horizontal axis relates to the number of data entries (time). In case of the comparison presented in Fig. 4 the log contained about 17000 data. The total number of different activities is 24 as presented in Fig. 3. That means that we use 24 spiking neurons. Thus, the 300 growing neurons and 24 spiking neurons together were able to memorize the pattern of the daily life log containing about 15000-30000 data.

5. Conclusion

In this paper growing neural gas and spiking neural network were applied to memorize the time series of life log information measured by smart phone sensors. The activity of the user can be estimated by the GNG and SNN

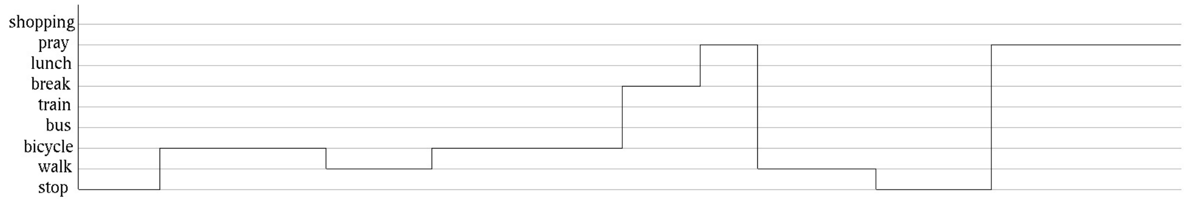

Fig. 5. Estimated data of subject 1

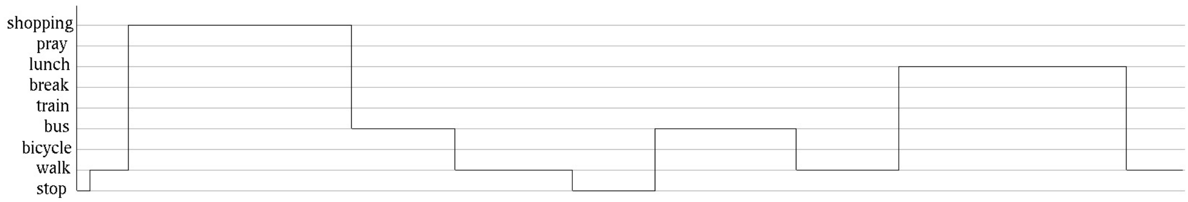

Fig. 6. Estimated data of subject 2 on the same day

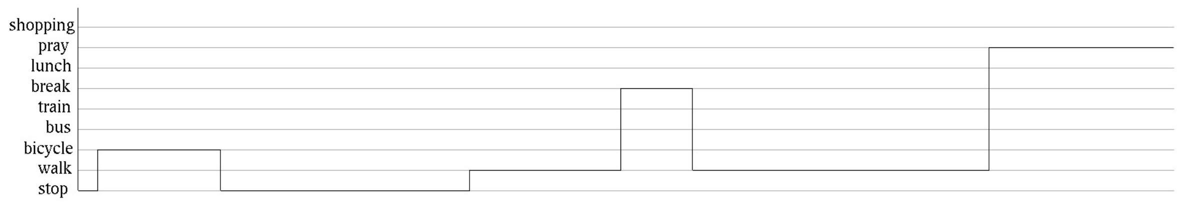

Fig. 7. Estimated data of subject 1 on the following day

Fig. 8. Estimated data of subject 3

techniques. The experimental results show the effectiveness of the neural computing by unsupervised learning.

One future work is to investigate different metrics in Eq. (22) and to apply not only the distance to the nearest growing neuron. Another aim is to create automatic analysis on the life log based on the extracted result by neural computing. A goal of the analysis is to understand cultural differences in the life log in order to support people with different cultural background.

Acknowledgements

This work was partially supported by MEXT Regional Innovation Strategy Support Program.

References

[1] C. M. Cordeiro, D. P. Agrawal, Ad Hoc & Sensor Networks – Theory and Applications, World Scientific Publishing, 2006.
[2] D. Tang, J. Botzheim, N. Kubota, T. Yamaguchi, Estimation of human transport modes by fuzzy spiking neural network and evolution strategy in informationally structured space, in: Proceedings of the IEEE Symposium Series on Computational Intelligence 2013, 6th IEEE International Workshop on Genetic and Evolutionary Fuzzy Systems, Singapore, 2013, pp. 36–43.
[3] J. Botzheim, N. Kubota, Growing neural gas for information extraction in gesture recognition and reproduction of robot partners, in: Proceedings of the 23rd International Symposium on Micro-NanoMechatronics and Human Science, Nagoya, Japan, 2012, pp. 149–154.
[4] N. Kubota, A. Yorita, Topological environment reconstruction in informationally structured space for pocket robot partners, in: Proceedings of the 2009 IEEE International Symposium on Computational Intelligence in Robotics and Automation, 2009, pp. 165–170.
[5] N. Kubota, D. Tang, T. Obo, S. Wakisaka, Localization of human based on fuzzy spiking neural network in informationally structured space, in: Proceedings of the IEEE World Congress on Computational Intelligence 2010, Barcelona, Spain, 2010, pp. 2209–2214.
[6] D. Tang, N. Kubota, Human localization by fuzzy spiking neural network based on informationally structured space, in: Proceedings of the 17th International Conference on Neural Information Processing, Sydney, Australia, 2010, pp. 25–32.
[7] D. Tang, N. Kubota, Information support system based on sensor networks, in: Proceedings of the World Automation Congress, Kobe, Japan, 2010.
[8] `http://www.apple.com`.
[9] T. Kohonen, Self-Organizing Maps, Springer-Verlag, 2001.
[10] T. Kohonen, Self-Organization and Associative Memory, Springer-Verlag, 1984.
[11] T. M. Martinetz, K. J. Schulten, A 'neural gas' network learns topologies, Artificial Neural Networks 1 (1991) 397–402.
[12] B. Fritzke, Unsupervised clustering with growing cell structures, Neural Networks 2 (1991) 531–536.
[13] B. Fritzke, Growing cell structures – a self-organizing network in *k* dimensions, Artificial Neural Networks 2 (2) (1994) 1051–1056.
[14] B. Fritzke, A growing neural gas network learns topologies, Advances in Neural Information Processing Systems 7 (1995) 625–632.
[15] B. Fritzke, Growing self-organizing networks – why?, in: Proceedings of the European Symposium on Artificial Neural Networks, 1996, pp. 61–72.
[16] W. Gerstner, W. M. Kistler, Spiking Neuron Models, Cambridge University Press, 2002.
[17] R. Hecht-Nielsen, Neurocomputing, Addison-Wesley, 1990.
[18] W. Maass, C. M. Bishop, Pulsed Neural Networks, MIT Press, 1999.
[19] J. M. Zurada, Introduction to Artificial Neural Systems, West Publishing Co., St. Paul, 1992.
[20] D. O. Hebb, The Organization of Behavior, Wiley & Sons, New York, 1949.