Classification of Trajectories Using Category Maps and U-Matrix to Predict Interests Used for Event Sites

Hirokazu Madokoro*, Kazuhito Sato, Nobuhiro Shimo
Faculty of Systems Science and Technology, Akita Prefectural University, 84-4, Tsuchiya Aza Ebinokuchi, Yurihonjo City, Akita, 015-0055, Japan

Abstract This paper presents a method for classification and recognition of behavior patterns based on interest from human trajectories at an event site. Our method creates models using Hidden Markov Models (HMMs) for each human trajectory quantized using One-Dimensional Self-Organizing Maps (1D-SOMs). Subsequently, we apply Two-Dimensional SOMs (2D-SOMs) for unsupervised classification of behavior patterns from features according to the distance between models. Furthermore, we use a Unified distance Matrix (U-Matrix) for visualizing category boundaries based on the Euclidean distance between weights of 2D-SOMs. Our method extracts typical behavior patterns and specific behavior patterns based on interest as ascertained using questionnaires. Then our method visualizes relations between these patterns. We regard our method as useful to estimate interest from behavior patterns at an event site.

Keywords Trajectory Analysis, Self-Organizing Maps, Hidden Markov Models, U-Matrix, Cross Validation

1 Introduction

Numerous security cameras are installed at public facilities such as stations, airports, schools, and hospitals for the purpose of preventing accidents and trouble of various kinds. Moreover, webcams that are easily accessed via the internet are widely used at kindergartens, nursing facilities, etc. to show the status of a family member, such as children or grandparents. Our society is therefore a so-called camera society. We spend our daily life in safety and with peace of mind because we are able to connect with each other via cameras irrespective of physical distances separating us.

As popularization of these consumer cameras continues, various services using camera images can be regarded as new life necessities. One such service is estimating the interest of a person using human trajectories obtained using image sequences obtained from cameras installed at amusement areas, shopping areas, event sites, etc. We can create a new service to present feedback of an exhibition or a booth in real time if we can estimate interest from human trajectories. Moreover, we can provide effective information to a host or organizers in real time.

Human trajectory analytical techniques obtained using a camera are used for marketing analyses [1], abnormal behavior detection [2, 3, 4], etc. For marketing analyses at large-scale stores such as department stores or small-scale shops such as convenience stores, trajectories of customers are analyzed for optimizing the layout of a shop or for reviewing showpieces. Moreover, we consider that the displayed interest is prone to affect human behavior and that it is identifiable by observation of that behavior because a product that is taken from a shelf or its neighbor products are often purchased. With respect to abnormal human behavior detection, trajectories are used to detect deviant behavior patterns that take different behavior patterns surrounding the person such as intrusions into restricted areas, running against a one-way street, and rapid turns [3, 4]. A suspicious person shows diverse behavior patterns when loitering or diverting attention with a group of people at a store or a station where numerous people congregate. We consider that human behavior patterns comprise two types: typical behavior patterns and specific behavior patterns.

The goal of this study is to classify behavior patterns based on interest from human trajectories at an event site. For trajectory analyses, it is a challenging task to extract a rule from numerous trajectory datasets automatically. According to the progress of the advanced performance of computers, machine-learning based methods that can extract rules automatically from given datasets through learning have been specifically examined recently. Machine learning is roughly classifiable into two types: supervised learning and unsuper-
As a study to classify human behavior patterns, Suzuki et al. [2] proposed a method to classify behavior patterns associated with purchasing. They detected deviation behavior patterns at retail stores using k-means for calculating probabilistic distance features between human trajectories using HMMs. However, k-means entails the number of classified behavior patterns in advance. Toyoshima et al. [3] proposed a method to discriminate lost behavior patterns or considering purchase using basic moving vectors that are calculated from basic behavior patterns such as walking, stop, and changing direction. For their method, the target behavior is of three patterns: lost behavior based on a scenario, purchasing behavior, and other free behavior. However, each pattern is fixed for a short time: only 90 s. Moreover, only three subjects were available for the experiment, and its environment remained as a virtual store, not an actual environment. They obtained a rule for analyzing behavior patterns from their simulation results obtained using decision trees. However, they discussed that this rule has no evidence for use as a statistical factor because of the paucity of data.

As a study to detect unsteady human behavior patterns, Suto et al. [4] used One Class Support Vector Machines (OC-SVMs) for learning features of central trajectories of dynamic object regions. They measured a distance from the decision boundary in a high-dimensional feature space as an index of non-stationarity. Poriki et al. [10] applied spectral clustering methods to HMMs. They detected unsteady events to classify behavior patterns of vector inner products. Nanri et al. [11] proposed a method to construct linear subspaces based on unsupervised learning. Using the measurement of unsteady of the distance between subspaces, they automatically extracted fallen pedestrians from video images. McKenna et al. [12] proposed a method using Expectation Maximization (EM) algorithms based on probabilistic models. From indoor video images, they estimated trajectories that a person might take when moving among chairs, a bed, and doorways. They judged the degree of stationarity of input samples probabilistically. Okamoto et al. [13] specifically examined differences between behavior patterns that recognize normal behavior according to circumstances in public spaces. They realized classification of behavior patterns for modeling human trajectories using HMMs in situations of departure or arrival of a train at a platform.

As a study to obtain trajectory datasets over a long period, especially for single life of a senior person, Aoki et al. [14] classified trajectory patterns to describe symbolic sequences of situations that were observed in daily activities of a target senior person. They detected extraordinary behavior to learn behavior patterns of attentional time zones using HMMs. However, the detection accuracy drops dramatically if extraordinary behavior patterns are included in the training phase. Shinagawa et al. [15] proposed a method to classify behavior patterns using Dynamic Planning (DP). They detected extraordinary days that differ from ordinary days of a senior person going out a long time, staying out overnight, or a guest staying at home, etc. However, they did not consider behavior patterns as representing bad health conditions of the target person. Moreover, future works are expected to develop a method for checking health status.

Numerous studies have been proposed for detecting unsteady behavior patterns. Nevertheless, these studies classify behavior patterns to two categories: steady or unsteady behavior patterns. As described in this paper, our method extracts typical behavior patterns and specific behavior patterns simultaneously at an event site that is estimated to have various behavior patterns. We attempt to visualize relations between trajectories for classifying behavior patterns based on unsupervised learning. Moreover, we recognize behavior patterns based on interest for remodeling trajectories using only extracted typical behavior patterns.

3 Proposed method

Figure 1 presents a flowchart of our method comprising the following six steps:

1. Vector quantization of trajectories using 1D-SOMs
2. Modeling trajectories using HMMs
3. Calculation of distance between HMMs
4. Behavior pattern classification using 2D-SOMs
5. Extraction of category boundaries using U-Matrix
6. Evaluation of recognition using Cross Validation (CV)

Steps 1)–3) correspond to preprocessing for extraction of features from trajectory datasets. Steps 4)–5) correspond to the main procedure of our method for behavior pattern classification based on unsupervised learning. In Step 6), we evaluate the recognition accuracy of behavior patterns after reconstructing models using classification results. The following are details of respective steps. Herein, we had proposed this basic method in the proceeding [21]. For this paper, we have improved our method to review detail procedures.

3.1 Vector quantization of trajectories using 1D-SOMs

Vector quantization is a method to approximate vector data to a finite number of representative symbols. We use 1D-SOMs for quantizing trajectories to a finite number of symbolic sequences. The input features of 1D-SOMs comprise three-dimensional features: \( x \) and \( y \) coordinate values of the position of a target person. Moreover, we used velocity to portray the displacement between frames. Based on results of our preliminary experiment for vector quantization, we set the number of codebooks to 25 units. The features and algorithms of 1D-SOM are the following.

The SOMs proposed by Kohonen [5] are unsupervised and self-organizing neural networks that learn input features with the concept of competition and neighborhood learning. The feature map, also called a category map, reflected input features generated through the learning of SOMs. The network comprises two layers: an input layer and a mapping layer. Each mapping layer unit has weights among all input layer units. Learning was conducted to update weights. The SOM learning algorithm is the following.

1. \( w_{ij}(t) \) are weights from an input layer unit \( i \) to a mapping layer unit \( j \) at time \( t \). These weights are initialized randomly.

2. Training data \( x_i(t) \) present input layer units \( i \) at time \( t \). The Euclidean distance \( d_j \) separating \( x_i(t) \) and \( w_{ij}(t) \) is calculated as

\[
d_j = \left( \sum_{i=1}^{I} (x_i(t) - w_{ij}(t))^2 \right)^{1/2}.
\]

3. The unit for which \( d_j \) is the smallest is defined as the winner unit \( c \).

4. \( w_{ij}(t) \) of \( c \) and its neighborhood units inside \( N_c(t) \) are updated as shown below.

\[
w_{ij}(t + 1) = w_{ij}(t) + \alpha(t)h_{ij}(t)(x_i(t) - w_{ij}(t))
\]

Therein, \( \alpha(t) \) is the learning rate coefficient and \( \sigma^2(t) \) is a function to adjust the size of the neighborhood region that decreases with the progress of learning.

5. The neighborhood region \( h_{ij}(t) \) around the winner units \( c \) is defined as

\[
h_{ij}(t) = \exp\left( -\frac{\|i - j\|^2}{2\sigma^2(t)}\right).
\]

6. Repeat 2)–5) up to the learning iteration that was set previously. For this study, we set the maximum number of iterations to 100,000 steps based on the results of a preliminary experiment.

3.2 Modeling trajectories using HMMs

For modeling human behavior patterns, HMMs are widely used in existing studies [2, 10, 13, 14] to contain robustness against displacement, expansion, and contraction of data. Actually, HMMs are of two types: discrete HMMs and continuous HMMs. Discrete HMMs are used if the distribution of features follows a normal distribution. Continuous HMMs are used if the distribution shape is unknown. For this study, we use discrete HMMs because various distributed shapes are presented for behavior patterns.

HMMs have several states. The emission probability of symbols in each state on HMMs is added to simple Markov models that are defined as transition probabilities between states. The learning of HMMs corresponds to estimates of two parameters: the status transition probability that produces training symbolic
sequences and symbolic emission probability. In the training phase, iterations are continued until the convergence of parameters becomes sufficiently small for emission likelihood variation using Baum–Welch algorithms.

For an HMM structure, we use ergodic HMMs that can learn several patterns from a single model. Unlike left-to-right HMMs, ergodic HMMs can return to the former state after transition to the next state. Figure 2 presents the network architecture of HMMs obtained using our method. The transition probability from state \( S_i \) to state \( S_j \) is depicted as \( a_{ij} \). The emission probability of symbol \( k \) at state \( S_i \) is depicted as \( b_i \).

### 3.3 Calculation of distance between HMMs

As a method for comparison of HMM models, Juang et al. [17] proposed a scale of the distance between HMMs using the Kullback information criterion. This scale is represented as the ratio of likelihood for the symbolic sequences of both models. The distance \( D \) between HMMs is defined as

\[
D(\lambda_1, \lambda_2) = H(X_1, \lambda_1) - H(X_1, \lambda_2),
\]

where \( \lambda_1 \) and \( \lambda_2 \) respectively portray the coordinate vector on the two-dimensional space of the winner unit and neighborhood units. As a result of our preliminary experiment, we allocated \( 14 \times 14 \) units to the mapping layer.

### 3.5 Extraction of category boundaries using the U-Matrix

Actually, U-Matrix is a method to visualize spatial distribution based on the distance of weights between units of SOMs. We measured the similarity between neighborhood units using a U-Matrix. Regarding labeling of SOMs, burst units for training data are assigned unique labels. Unburst units must be assigned labels used for testing data. Our method creates clusters of units that maintain similar weights from category boundaries using U-Matrix. This procedure must be performed manually if we do not use a U-Matrix. We consider that using a U-Matrix has benefits for creating clusters for labeling.

The brief algorithm used to extract boundaries on the U-Matrix is presented as the following. The unit \((j, k)\), which is allocated a rectangular pattern, contains eight neighborhood units depicted in Fig. 3. The U-matrix value \( U_{\text{mat}} \) is calculated as the degree of similarity. For neighborhood units between the vertical and horizontal directions, \( U_{\text{mat}} \) between \((j, k)\) and \((j+1, k)\) is calculated as

\[
U_{\text{mat}}((j, k)(j+1, k)) = |w(j + 1, k) - w(j, k)|
\]

where \( w(j, k) \) is a weight in the grid \((j, k)\). For the neighborhood units of the oblique directions, \( U_{\text{mat}} \) between \((j, k)\) and \((j+1, k+1)\) is calculated as

\[
U_{\text{mat}}((j, k)(j+1, k+1)) = \frac{1}{2}(|w(j, k) - w(j + 1, k)| + |w(j, k + 1) - w(j + 1, k)|).
\]
4.2 Questionnaire

For this study, we took a questionnaire for identifying the interest of a visitor to a booth. The questionnaire comprises three terms: attributions such as age, interesting booths, and the satisfying degree for exhibition. The term of interesting booths can check multiples because it is not always one booth to have interest by a visitor. For considering user-friendliness of selection, four levels can be selected for satisfying the degree for exhibition: strong pleasure, pleasure, displeasure, and strong displeasure.

The procedures for visitors are the following:

1. to receive a questionnaire,
2. enter the room from the right-side door,
3. watch booths freely,
4. exit to the left-side door,
5. to write a questionnaire.

We obtained approval from a subject through prior oral explanation. Nevertheless, we did not explain the purpose of this study, thereby preventing effects of intentional behavior.

4.3 Trajectories

The field examination was conducted at an Open Campus held on 16 July 2010 at our university. During events lasting seven hours from 9:00AM through 4:00PM, 41 people visited our laboratory. The age distribution is 60 % in their 20s or younger, and 40 % in their 30s or older. Table 1 presents details related to the number of visitors and their interesting booths. The combinations of interesting booths were seven types: RB, FB, MB, RB & FB, RB & MB, FB & MB, and

used our laboratory as an event site for this experiment. Figure 4 presents the layout of our laboratory. The floor space is 98 m$^2$, 7 m long by 14 m wide, forming a rectangular space.

The exhibition booths based on our research contents are three, the Robot Booth (RB), Facial Booth (FB), and Medical Booth (MB). The room has two doors: The upper-right side door is the entrance; the upper-left side door is the exit, as depicted in Fig. 4.

We obtained time-series images using two network cameras (VB-C500VD; Canon Inc.). The camera angle of the cameras is 82.0 deg horizontal and 60.5 deg vertical. The black marks and dashed lines respectively represent the installed positions and the view range of the cameras. The two cameras were synchronized automatically via the network. The bottom part of Fig. 4 is beyond the view range of the cameras. Therefore, we installed posters related to exhibition to constrain visitors’ moving inside the view range.

Figure 5 presents an image of the room obtained using the cameras. The camera resolution is 640 × 480 pixel. Regarding a processing cost and a disk space, we converted original images for down-sampling from 5 fps to 1 fps.

### 3.6 Evaluation of recognition with CV

We create models in each behavior pattern classified using 2D-SOMs and U-Matrix. We evaluate the recognition accuracy of behavior patterns to compare the likelihoods of respective models. For creating models, we use HMMs similarly for modeling trajectories described in the former section. As a classifier used for HMMs, we create models based on supervised learning. We use leave-one-out CV [18] to evaluate models. We remove one trajectory to create a training dataset. Subsequently, we evaluate the generalization capability for the removed trajectory. Here, the recognition accuracy RA is defined as

$$RA = \frac{G \cap I}{N} \times 100$$

(12)

where $N$ signifies the number of visitors, $G$ denotes ground truth data, and $I$ stands for an index of a HMM model that shows the maximum likelihood.

### 4 Datasets

#### 4.1 Environment

For existing studies of trajectory analyses, no dataset has been standardized sufficiently. Therefore, originally developed datasets are used [1, 2, 3, 4]. For this study, we created an original dataset for classification of behavior patterns based on interest at an event site. We...
Table 1. Interested booths and the numbers of visitors.

| Interested booth     | Number of visitors |
|----------------------|--------------------|
| Robot Booth (RB)     | 10                 |
| Facial Booth (FB)    | 3                  |
| Medical Booth (MB)   | 1                  |
| RB & FB              | 8                  |
| RB & MB              | 4                  |
| FB & MB              | 2                  |
| All Booths (AB)      | 13                 |
| Total                | 41                 |

All Booths (AB). Nevertheless, the distribution of the number of visitors was uneven.

As portrayed in Table 1, the visitors were fewer than four persons in FB, MB, RB&MB, and FB&MB. The approximate accuracy of probabilistic distribution on HMMs drops dramatically if the training data are insufficient for the modeling phase [16]. Moreover, it is a challenging task for HMMs to learn and to produce a model of behavior patterns from characteristic patterns. Therefore, we used trajectories of RB, RB&FB, and AB such that the data are sufficient compared with other trajectories. The total data include 31 trajectories.

We developed a method for extracting trajectories from time-series images based on results of an earlier study [19]. We manually corrected frames for which positional coordinates of a target person could not be extracted automatically because of occlusion or corruption. Regarding positional coordinates obtained from time-series images, we extracted three-dimensional features: x and y coordinates and velocity.

5 Experimental results

At the beginning of the experiment, we optimized HMM states as the most important parameters for modeling trajectories. Subsequently, we extracted typical behavior patterns and specific behavior patterns based on interest after classifying the trajectories. Moreover, we created models in each interesting booth using trajectories that were removed specific behavior patterns. Finally, we evaluated the recognition accuracy of behavior patterns in each model based on leave-one-out CV.

5.1 Optimization of HMM parameters

The HMM performance depends strongly on the structure that is defined by the number of states and their transitions [20]. As the definition of HMM states, Suzuki et al. [2] used segmented regions related to actual spaces instead of the information source that is used generally. Similarly, we allocated HMM states to each region to be segmented in the room of our laboratory.

Figure 6 presents recognition accuracy to change the number of HMM states from step 4 through 16 by two and the detected number of NaN.

5.2 Behavior pattern classification

Figure 7 presents a classification result of trajectories for a category map using 2D-SOMs and the U-Matrix. Labels 1–3 respectively correspond to RB, RB&FB, and AB. Filled black regions depict category boundaries of the U-Matrix. Label 1 is distributed in the upper and lower-right regions on the category map. Label 2 is distributed in the lower-left, central, and upper-right regions. Label 3 is distributed in the central and left regions. For this result, trajectories are classified globally in each booth according to interest. However, no category is formulated according to behavior patterns with the same interest as specifically examining the regions to be divided by boundaries. This result means that behavior patterns classified using 2D-SOM-based signal features do not always correspond to behavior patterns intentionally labeled as obtained using questionnaires.

Behavior patterns of those interested in RB are broadly divisible into two categories because two clusters of Label 1 are apparent on the category map. In Label 2, trajectories are classified to the lower-left part on the category map. However, one trajectory is apparent in the center part. Moreover, one trajectory is apparent in the upper-right part. We respectively marked these units in the center and upper-right with Rectangles A and B. The human trajectories are shown respectively in Figs. 8 and 9. White lines depict trajectories from the entrance door to the exit door of the room. The trajectories depicted in Figs. 8 and 9 respectively comprise 1,198 frames (19 m 58 s) and 558 frames (9 m 18 s). In the division with category boundaries, the cluster
mapped Label 2 in Rectangle A comprises five trajectories of Label 3 and one trajectory of Label 2.

Label 2 corresponds to trajectories that are interested in RB and FB obtained from questionnaires. In the trajectory shown in Fig. 8, the subject spent a long time at RB and FB. The subject stayed in front of a poster that had been put up between the bookshelf and the exit door. He did not access the booth around the desk where a demonstration of facial expression studies was held. In the questionnaire, the subject answered that he is interested in FB. However, he did not access FB, practically. Therefore, we regard this trajectory as classified to typical data except for the behavior pattern of other data in Label 2.

Rectangle C portrays a trajectory of Label 3 located at the center on the category map near the Rectangle A. Figure 10 depicts a trajectory of Rectangle C. This trajectory comprises 1,181 frames (19 m 41 s). Results of the questionnaire reveal that the subject marked Rectangle A as a typical behavior pattern. His behavior pattern resembles the behavior pattern of AB on the category map. We regard the trajectory of Rectangle A as a specific behavior pattern of Label 2, which is classified to the lower-left part in Fig. 7, which is a typical behavior pattern of RB&FB.

Figure 11 depicts a trajectory of Rectangle D as a representative example of a typical behavior pattern of Label 2. Rectangle D contains a trajectory that is located to the center of Label 2 to be classified to the lower-left part. The trajectory comprises 1,051 frames (17 m 31 s). As a machine-learning based method to extract a rule from presented data automatically, it is a challenging task to specify a feature as a rule. While observing trajectories visually, it is apparent that the feature that a subject stays at each booth and moves around there in Figs. 8 and 11. The difference of staying patterns at FB is observed in the center of the image shown in Figs. 10 and 9. Moreover, the tendency to move subjects straight to the exit door except for those staying at RB is observed on the right side of the image shown in Fig. 9. We regard these patterns as represented on the category map.

As an existing method to detect unsteady behavior patterns, Sutou et al. [4] used OC-SVMs for detecting outlier trajectories. Moreover, HMMs are used for determining a threshold for a likelihood [2, 13]. In our method, specific behavior patterns can be detected for visualizing the relation between trajectories using 2D-SOMs and U-Matrix. Moreover, behavior patterns similar to a typical behavior pattern can be detected visually through the category map.

As the number of categories, our method classifies trajectories without the previous setting of the number of categories because of employment of SOMs. We extracted category boundaries using U-Matrix. However, we have not obtained a result to extract a suitable number of categories. In our datasets, we considered that the suitable number of categories classified automatically is three patterns. However, confusions of labels are respectively apparent in different or similar feature patterns among the same and different categories.

Actually, U-Matrix is extremely useful for feature classification of trajectories. However, our results have no consistency between subjective labels from questionnaires. A tendency of category creation is observed partially on the units assigned Label 3 around the center of the category map. In Fig. 10, Rectangle C depicts a typical pattern for the feature of this trajectory. As a visual evaluation of trajectory features, this subject stayed at all booths for a fixed term. Moreover, for the trajectory marked Rectangle B in Fig. 9 as a typical behavior pattern of Label 2, the subject stayed at the right in RB and the center in FB. However, staying at the left in MB is not observed. We regard this as a challenging task to analyze the parts of local features directly using SOMs because these trajectories were converted features using HMMs. In the next section, we evaluated classification results obtained using a supervised learning-based method as recognition accuracy.

5.3 Behavior pattern recognition

Table 2. Recognition accuracy of each model.

|        | RB    | RB&FB | AB    | Ave.  |
|--------|-------|-------|-------|-------|
| Before | 10.0% | 50.0% | 53.8% | 38.7% |
| After  | 10.0% | 83.3% | 61.5% | 48.5% |

It is a challenging task to learn and model features from a mixed dataset of specific and typical behavior patterns. We consider that recognition accuracy might be improved by detecting specific behavior patterns and removing them from the dataset. Herein, the boundary information is obtained using the U-Matrix for extracting specific behavior patterns. However, we have not realized an automatic boundary extraction method using U-Matrix yet. Therefore, we extracted those related data from the category map based on our experience.

In this experiment, we evaluated models of two types.
before and after behavior pattern classification. For the models after behavior pattern classification, we removed trajectories that were classified as having specific behavior patterns. Specifically, we removed two trajectories of Rectangles A and B shown in Fig. 7, extracted as specific behavior patterns of RB&FB. We did not specify them as specific patterns because the same labels exist in neighboring areas, not Rectangles A or B, although a part without separated Labels 2 or 3 exists in the left-lower side of the category map shown in Fig. 7.

We evaluated each model based on leave-one-out CV to calculate the likelihood in each model. We determined that behavior patterns are recognized if a model whose likelihood is the highest corresponded to a booth in which a person is interested. For numerical evaluation of behavior pattern recognition, we calculated the recognition accuracy in each model.

Table 2 depicts the recognition accuracy of models before and after behavior pattern classification. The recognition accuracy values before behavior pattern classification are, respectively, 10.0%, 50.0%, and 53.8% for RB, RB&FB, and AB. The mean recognition accuracy is 38.7%. The recognition accuracy values obtained after behavior pattern classification are, respectively, 10.0%, 83.3%, and 61.5% for RB, RB&FB, and AB. The mean recognition accuracy is 48.3%.

The recognition accuracy before and after classification of RB remains at about 10.0%. This result is lower than that of other booths. As described in the results presented in the previous section, we consider that RB includes behavior patterns of two types. In this experiment, we used a single HMM model. We consider that classification accuracy can be improved if we use two HMM models in each behavior pattern. However, for this experiment, we divided no models since NaN might be increased because they lack the number of training data.

The recognition accuracy of RB&FB is improved by 33.0 percentage points: from 50.0% to 83.3% by behavior pattern classification. Moreover, the recognition accuracy of AB is improved 9.6% from 38.7%–61.5%. We consider that we can obtain a result to estimate interest from behavior patterns.

6 Conclusion

This paper presented a method for classifying behavior patterns based on unsupervised learning aimed at behavior pattern recognition related to interest inferred from human trajectories at an event site. Our method extracted typical and specific behavior patterns with consideration of similarity of trajectories for visualizing relations between behavior patterns using 2D-SOMs and the U-Matrix. Moreover, we evaluated our method based on leave-one-out CV after remodeling trajectories using typical behavior patterns. Comparing the performance measured before and after behavior pattern classification, the recognition accuracy was improved by
9.6 percentage points. We infer that we can obtain a result demonstrating an estimation of interest based on people’s behavior patterns.

Future work includes the use of automatic parameter optimization such as the number of codebooks for vector quantization, the number of mapping layer units on SOMs, and the threshold value of categories on U-Matrix to be changed according to trajectories. Moreover, we will improve the recognition accuracy of HMMs for application of our method to various environments. Furthermore, we will apply Latent Conditional Random Fields for modeling the human behavior patterns.

REFERENCES

[1] T. Yoshida, T. Koisho, K. Hattori, I. Toyoshima, and N. Imasaki, “A Framework for Improving Profitability in a Retail Store Using Customer Behavior Analysis,” IPSJ SIG Notes, vol. 24, pp. 121–125, Mar. 2005.

[2] N. Suzuki, K. Hirasawa, K. Tanaka, Y. Kobayashi, Y. Sato, and Y. Fujino, “Detection of Abnormal Behavior and Patterns by Human Trajectories Analysis,” IEICE Trans. on Information and Systems, vol. J91-D, no. 6, pp. 1550–1560, June 2008.

[3] I. Toyoshima, T. Itakura, K. Hattori, T. Yoshida, and T. Koiso, “A Study of classification of human behavior using walking trajectory data,” Technical report of IEICE. CST, vol. 106, no. 180, pp. 19–24, Jul. 2006.

[4] K. Sudo, K. Wakabayashi, K. Arakawa, and T. Yasuno, “Detecting Anomalous Sequences in Long Duration Monitoring Videos,” IPSJ SIG Notes. CVIM, vol. 112, pp. 77–82, Nov. 2005.

[5] T. Kohonen, Self-Organizing Maps, Springer Series in Information Sciences, 1995.

[6] S. Takahashi, “Speech recognition using HMMs: The ability of representation and the robustness in recognition,” IEICE technical report. Speech, vol. 93, no. 32, pp. 1–6, May 1993.

[7] A. Ultsch, “Clustering with SOM U C,” Proc. Workshop on Self-Organizing Maps ’05, pp. 75–82, 2005.

[8] C. Cortes and V. Vapnik, “Support-Vector Networks,” Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.

[9] J. McQueen, “Some Methods for Classification and Analysis of Multivariate Observations,” Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability, pp. 281–297, 1967.

[10] F. Porikli and T. Haga, “Event detection by eigenvector decomposition using object and frame features,” Proc. Computer Vision and Pattern Recognition, pp. 114–114, June 2004.

[11] T. Nanri and N. Otsu, “Unsupervised Abnormality Detection in Video Surveillance,” Proc. IAPR Conference on Machine Vision Application, pp. 574–577, 2005.

[12] S.J. McKenna and H. Nait-Charif, “Learning Spatial Context from Tracking using Penalised Likelihoods,”
H. Okamoto, S. Nishio, N. Babaguchi, F. Morii, and N. Hagita, “Situation-based behavior modeling for anomaly detection,” Technical report of IEICE. PRMU, vol. 108, no. 94, pp. 103-108, June 2008.

S. Aoki, M. Onishi, A. Kojima, and K. Fukunaga, “Detection of a Solitary Senior’s Irregular States Based on Learning and Recognizing of Behavioral Patterns,” IEEJ Transactions on Sensors and Micromachines, vol. 125, no. 6, pp.259–265, June 2005.

Y. Shinagawa, T. Kishimoto, and S. Ohta, “Detection of an Unusual Day for the Elderly Living Alone Using the Classification of Movement Pattern,” Kawasaki Medical Welfare Journal, vol.15, no. 1, pp. 175–181, 2005.

L. Zhang and Q. Ji, “Image Segmentation with a Unified Graphical Model,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 8, pp. 1406–1425, July 2010.

B.H. Juang and L.R. Rabiner, “A probabilistic distance measure for hidden Markov models,” AT&T Tech. J, vol. 64, no. 2, pp. 391–408, Feb. 1985.

C.M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.

R. Matsumura and K. Okamura, “Object Detection and Tracking Using Particle Filter,” Annals of Oshima National College of Maritime Technology, vol. 41, pp. 75–85, Dec. 2008.

S. Ikeda, “Construction of Phone HMM Using Model Search Method,” Trans. of IEICE, vol. J78-D-2, no. 1, pp. 10–18, Jan. 1995.

H. Madokoro, K. Honma, and K. Sato, “Classification of Behavior Patterns with Trajectory Analysis Used for Event Site,” Proc. IEEE International Joint Conference on Neural Networks, pp.1449-1456, June. 2012.