Service-Operation Estimation in a Japanese Restaurant Using Multi-Sensor and POS Data

Ryuhei Tenmoku[1], Ryoko Ueoka[1], Koji Makita[1], Takeshi Shinmura[1,2], Masanori Takehara[3], Satoshi Tamura[3], Satoru Hayamizu[3], and Takeshi Kurata[1]

[1] National Institute of Advanced Industrial Science and Technology, Central 2, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan
r-tenmoku@aist.go.jp
[2] Ganko Food Service co. ltd.
2-13 Shinkitano 1-chome, Yodogawa-ku, Osaka 532-0025, Japan
[3] Gifu University.
1-1 Yanagido, Gifu, 501-1193 Japan

Abstract. This paper describes the strategy of our SOE (Service Operation Estimation) which is the operational behavior estimation of employees in labor-intensive service fields. We assume that SOE would be utilized for the feedback for employers and employees in all kinds of the service field: service operational improvement, employee training, and QC (Quality Control) activities. In this paper, we introduce our SOE framework and a case example of our SOE of serving staffs in a Japanese restaurant using multi-sensor data and POS data.

Keywords: Service-Operation Estimation (SOE), Action Recognition, Voice Activity Detection (VAD), Machine Learning, Service Engineering

1 Introduction

Improvement of productivity in service industries plays an important part of continuous economic growth. Service engineering, the framework which erects scientific approaches to achieve improvement of productivity in service industries receives much attention as a new field of engineering [1]. Authors consider that it is important that to realize the “OADI cycle” (observation, analysis, design, and implementation) in the service industry makes a large contribution to improvement of productivity [2]. 4 Steps of the OADI cycle is as follows:

1. Observation: Observing actions of service providers and receivers (customers) in actual service fields
2. Analysis: Analyzing obtained data
3. Design: Designing models based on objective evidences obtained in the previous step
4. Implementation: Implementing designed models in service fields

Authors have realized a feedback system of analyzed and visualized results of employees' positions on duty. We assume that such kind of information can be available for service operational improvement, employee training, and QC (Quality Control) activities. As the next step of this feedback, visualizing employees' operational behaviors as they're working is demanded. If we obtained service operations (SOs), we could realize operation analysis with employees' physical load, employee training considering personal behavior patterns, and supporting description of work journals and handover documents.

We can consider that employees' SOs seem to have a correlation with employees' behavior data, work-related data, and accounting data (called them "components" in this paper). Based on this idea, we construct a framework of estimating employees' SOs from these components using machine learning techniques (especially, supervised learning including a k-nearest neighbor method [3], a support vector machine [4], and some boosting methods [5]). In this paper, as one of implementation of practical SOE systems, we introduce SOE of serving staffs in a Japanese restaurant using multi-sensor and POS (Point of Sales) data.

2 SOE Framework

2.1 Related Works

In wearable and ubiquitous computing societies, many researchers try to recognize human behaviors. These researches can be categorized into sensor-based methods, image-based methods, and hybrid methods of them. However, most of these works aim to estimate human primitive actions (running, walking, standing up, sitting down, and so on) [6] or daily activities (drinking coffee, watching TV, washing face, and so on) [7].

Some researches aim to estimate operational behaviors of workers. Hartmann proposed a labors' behavior estimation method in an auto factory using a hybrid method [8]. Chae proposed a workers' safety management system in construction sites [9]. Chae also proposed a behavior estimation method of construction labors. Ohmura proposed a semi-automatic behavior recording system of nurses [10].

2.2 Outline of SOE Framework

Again, our goal is realizing SOE which can be used in various kinds of service fields in the long term. In this scenario, we have to fulfill the following requirements:

- It is easy to apply our SOE system for every kind of the employee's role and service field. Obviously, SOs, components, and the relationship of them vary by employee's role and service field.
- We can use components which can be obtained without employees' overload and invading customers' privacies. In principle, we cannot use any cameras in areas where customers can move into.

To satisfy these demands, we chose the method which constructs SO classifiers peculiar to service fields and roles of employees using components which can be obtained easily.

In our SOE framework, behaviors of employees in the service fields are obtained by the off-line estimation from the following components.

(a) Behavior data: position, orientation, action, voice activity detection (VAD), and keyword spotting data

(b) Work-related data: shift data, attendance management data, operational schedule, and some digital logging operational information

(c) Accounting data: POS data and cash register data

We realize SOE from time-synchronized components by selecting the most probable SO from some pre-determined SO candidates according to the role of the employee. Before SOE, the user prepares components with true SO values to construct SO classifiers using supervised learning methods.

3 Case Example and Implementation

We have measured components of approximately 20 employees in a Japanese restaurant "Ganko Ginza 4-Chome" to realize their SOE. This restaurant consists of 2 floors and has several types of customer seats: counter seats, dining tables, private dining rooms, and party rooms. This measurement experiment was held for 4 consecutive weeks. In this restaurant, employees are categorized into the following roles: kitchen staff, bus boy, and serving staff. In this paper, we introduce case example of serving staffs' SOE.

3.1 Measurement of Components

Every serving staff equips two sensors: a PDR sensor on the back of the waist and a bone conduction microphone. PDR sensors can measure the employees' position and orientation and detect several primitive actions based on the movement of body trunk. The employee's position and orientation can be obtained based on our pedestrian tracking system for indoor environments [11]. This tracking system utilizes infrastructures such as surveillance cameras and active RFID tags which are sparsely put in place to correct the tracking errors of our pedestrian dead reckoning [12].

We extracted "area information" which was rough and meaningful information from employee's position data. In particular, we separated the restaurant into 7 areas (a kitchen, aisles, customer seats, pantries, cash registers, stairways, and an entrance) as shown in Fig.1 and employees' positions were expressed as such areas instead of 3D coordinate values. Orientation data was also presented as discrete values in a similar way. In the current system, we can detect horizontal and vertical actions (walking and standing up / sitting down).

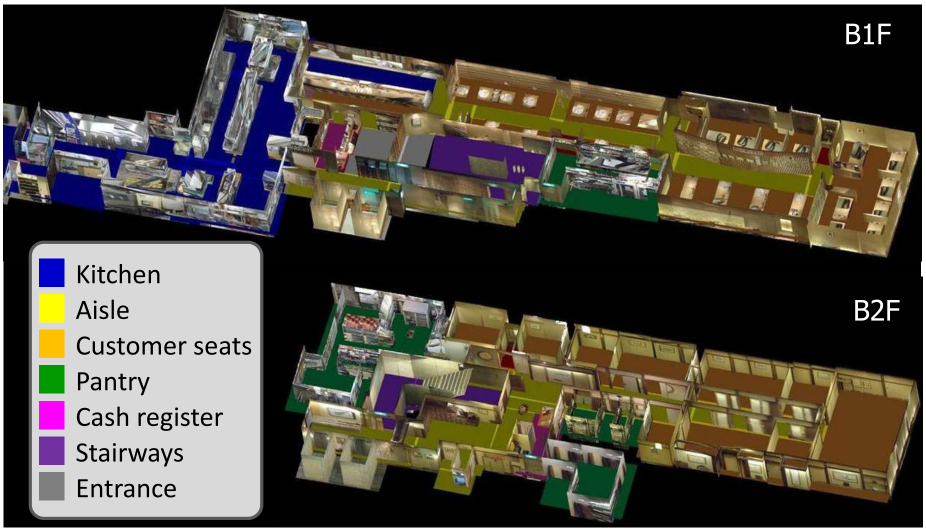

Fig.1 Areas definition in a Japanese restaurant

A bone conduction microphone records a serving staff's speech while he/she is working. In our current work, we detect only if the worker is speaking or not (VAD), however, we plan to detect some operational keywords from his/her speech data (keyword spotting). They are especially effective material to estimate the behavior of employees who serve customers. Of particular note is that using bone conduction microphones is effective to the privacy of customers since they can reduce every sound other than employees' voices. Additionally, we use POS data as an accounting data. The POS data which we used includes time, table numbers, handling staff names, and contents of all orders.

3.2 Implementation of Local SOE Using AdaBoost

Fig. 2 shows a block diagram of local SOE. We have 2 phases to realize SOE. In the first phase, learning phase, we prepares weak classifiers and components with true SO values (samples). True SO values are given manually as hearing voice data or watching video data. Some features are extracted from these components. In our implementation, 56 features are extracted from components and breakdown of these features are shown in Table1. Features assume to be extracted from single component or combined data of them, for example, time amount of VAD, detecting a certain action in a certain area, and so on. Weak classifiers are threshold functions of these features. In the AdaBoost framework, the strong classifier is constructed by choosing weak classifiers and determining their weighs based on the relationship of features and true SO values of samples. In the estimation phase, 56 features are extracted from components without true SO in the same way as the learning phase. In this phase, SO is estimated by applying the strong classifier which is constructed in the previous phase to features.

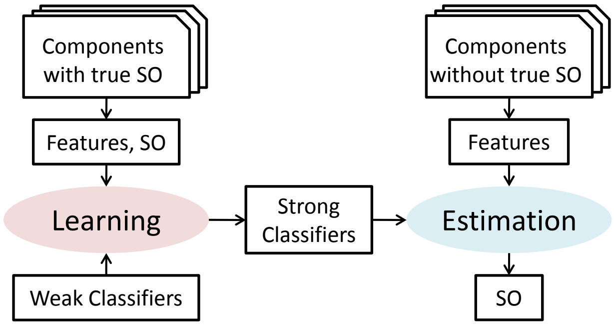

Fig.2 Block diagram of Local SOE

Table 1 Breakdown of used features

Position (Area)	7
VAD	1
Motion	2
Orientation	4
POS data	4
Combined features	38

The service operations of serving staffs are defined according to the demand of employers of this restaurant as follows:

(1) Taking orders
(2) Serving food or drink
(3) Moving and carrying something
(4) Accounting
(5) Greeting and offering customers to the table
(6) Cleaning up and setting tables
(7) Dialogue with customers
(8) Dialogue with other staff

Since AdaBoost can construct 2 classes classifiers, we construct 8 strong classifiers which can divide into each SO and the rest. In other words, outputs of these strong classifiers indicate likelihood values of each SO.

3.3 Integration of Results of Local SOE

In the previous section, we show the method which can estimate SOs of consecutive components. This local SOE has two parameters: the time and the length of time the local SOE is applied. We realize SOE including segmentation by integrating results of local SOE. In our implementation, the integrated result of SOE of the time t is defined as the SO which has the maximum integration score. Here, the integration score of each SO is defined as the sum of products of the weighting factor and score of local SO including the time t.

4 Experiments and Results

We had two kinds of experiments about serving staffs' SOE in the Japanese restaurant. First, we show each feature's degree of contribution in local SOE. Second, we show the accuracy rate of integrated results of our SOE.

Table 2 Contribution of each feature category
(Red numbers show the top 3 feature categories in each SO.)

Feature categories	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Area	12.5	12.1	6.7	33.5	22.3	5.4	13.9	8.9
VAD	6.8	11.2	2.0	8.3	15.5	7.3	11.7	7.6
Motion	4.7	6.4	7.3	48.7	18.6	10.0	11.7	5.8
Orientation	5.9	12.9	7.5	26.4	24.6	21.3	21.1	10.2
POS data	2.6	2.9	0.7	7.0	5.6	0.6	6.3	0.6
Area, VAD	19.7	2.8	4.2	38.9	23.9	6.7	6.2	7.5
Area, Motion	15.5	2.4	6.9	46.8	21.4	16.3	6.8	8.5
Area, Orientation	30.5	29.6	12.3	78.1	95.7	30.6	33.2	13.9
VAD, Motion	5.5	15.9	3.6	1.4	9.1	4.5	11.9	3.0
VAD, Orientation	17.3	20.1	6.8	1.3	63.3	9.2	8.6	10.5
Motion, Orientation	15.9	23.7	12.0	67.1	11.2	16.6	21.0	13.5

4.1 Contribution of Features in local SOE

We constructed strong classifiers of 8 kinds of SO discussed in section 3.2 using approximately 4 hours components with true SO values. Here, the number of AdaBoost round was 1,000 in each SO. Table 2 shows summations of weight values of weak classifiers which were chosen in each round by feature categories in each SO. The larger these values, the higher the contribution of the feature to strong classifiers of the SO (the higher the correlation between the feature and SO). However, since absolute values of them depend on the amount of sample data and variability of features, we cannot compare these values between different SOs. That is to say, we can compare two values in the same column, but not in the same row.

Table 2 tells us orientation features have the strongest correlation with each SO in features extracted from one component. Especially, combined features of orientation and area have so strong correlation with each SO.

4.2 Accuracy Rate of SOE

We had a following experiment to show the accuracy rate of integrated SOE. First, we had local SOE of components with true SO using strong classifiers which was constructed in Section 4.1. Here, the resolution of the time of local SOE was 1 second and we had 5 trials under the conditions that the length of time was 5, 10, 15, 20, and 25 seconds, respectively. Second, these results of local SOE are integrated into the conclusive results of SOE. We adopted inverse numbers of the time length as weighting factors. Under such conditions, the rate that the most probable SO corresponded to the true SO was 81% and the rate that the most or the second most probable corresponded to the true SO was 90%.

5 Summary

This paper describes the strategy of our SOE and shows a case example of our SOE of serving staffs in a Japanese restaurant. We also show some experiments and results to show the feasibility of our SOE. We think any information can be used for SOE as far as they have any correlation with true SOs and are connected to the time stamp. If we use many kinds of data correlated strongly with service operations, we can obtain accurate results of SOE. However, it is important to consider the balance between estimation accuracy and the measuring cost and load.

In the future, we will try to develop our SOE framework into a hybrid system of machine learning and rule-based method. If the user can choose the balance of them for constructing strong classifiers, we can reduce the initial cost of our SOE system. Additionally, we plan to improve our visualizing tool which can play back of synchronized components and SOE results, realizing SOE in various service fields, and QC activities / operational improvements using SOE results.

Acknowledgement

This study was partially supported by the project of promotion of Service Engineering from Japanese Ministry of Economy, Trade and Industry (METI) in 2010 and 2011.

References

1. T. Tomiyama: "A Manufacturing Paradigm toward the 21st Century," Computer Aided Engineering, Vol. 4, pp.159 - 178 (1997)
2. T. Kurata, M. Kourogu, T. Ishikawa, J. Hyun, and A. Park: "Service Cooperation and Co-creative Intelligence Cycle Based on Mixed-Reality Technology," Proc. the 8th Int. Conf. on Industrial Informatics, pp.967 - 972 (2010)
3. B. V. Dasarathy: "Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques," IEEE Computer Society Press (1990)
4. N. Cristianini and J. S. Taylor: "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods," Cambridge University Press (2000)
5. Y. Freund and R. E. Schapire: "A Decision-theoretic Generalization of On-line Learning and an Application to Boosting," Jour. of Computer and System Sciences, Vol. 55, No. 1, pp. 119 - 139 (1997)
6. S. S. Intille, L. B. Emmanuel, M. Tapia, and J. Rondoni: "Acquiring In Situ Training Data for Context-Aware Ubiquitous Computing Applications," Proc of CHI2004, Vol. 6, No. 1, pp. 1 - 8 (2004)
7. L. Wang, T. Gu, X. Tao, and J. Lu: "Sensor-Based Human Activity Recognition in a Multi-user Scenario," Computer Science, Vol. 5859, pp. 78 - 87 (2009)
8. B. Hartmann, C. Schauer, and N. Link: "Worker Behavior Interpretation for Flexible Production," Proc of Int. Conf. on Computer, Electrical, and Systems Science, and Engineering, Vol. 58, pp. 494 - 502 (2009)

9. S. Chae and T. Yoshida: “A Study of Safety Management Using Working Area Information on Construction Site,” Proc on the 25th Int. Symp. On Automation and Robotics in Construction, pp.292 - 299 (2008)
10. R. Ohmura, F. Naya, H. Noma, N. Kuwahara, T. Toriyama, and K. Kogure: “Practical Design of A Sensor Network for Understanding Nursing Activities,” Proc. of Conf. on Local Computer Networks, pp. 615 - 622 (2006)
11. T. Ishikawa, M. Kourogi and T. Kurata: “Economic and Synergistic Pedestrian Tracking System with Service Cooperation for Indoor Environments,” Int. Jour. of Organizational and Collective Intelligence, Vol. 2, No. 1 pp. 1 - 20 (2011)
12. M. Kourogi and T. Kurata: “Personal Positioning based on Walking Locomotion Analysis with Self-Contained Sensors and a Wearable Camera,” Proc. the 2nd Int. Symp. on Mixed and Augmented Reality (ISMAR03), pp. 103 - 112 (2003)