The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Estimation of operational intentions utilizing Self-Organizing Map with Bayes filtering

Satoshi Suzuki and Fumio Harashima

Abstract— **An estimation algorithm of operational intentions in the machine operation is presented in this paper. State transition relation of intentions was formed using Self-Organizing Map (SOM) from the measured data of the operation and environmental variables with the reference intention sequence. Operational intention was estimated by stochastic computation using a Bayesian particle filter with the trained SOM. The presented algorithm was applied to the remote operational task, and qualitative and quantitative analyses were performed. As a result, it was confirmed that the estimator could classify the types of intentions as similarly as the human analyst discerned. Further, several issues, such as difficulty in preparation of objective normative data, and necessity of consideration of scenario / causality, are discussed.**

I. Introduction

Estimation of human intention is quite practical for various applications such as marketing and designs of man-machine interfaces. Recently, development of various clustering methods such as a support-vector-machine and a k-nearest neighbor method make it possible to realize applications of the intention estimation such as a web search engine and the customer trend analysis. There are many other systems that will be enhanced if such estimation techniques is embedded into them. A human-machine system like vehicles that human operator manipulates is one of them. In such system, not only improvement of the safety and efficiency but also enhancement of driver's ability can be expected [2]. Estimation of intentions in the human-machine system is, however, more difficult than the successful examples mentioned above. The followings are the reasons.

I1) It is difficult to identify the information types which are utilized for the decision making.

I2) It is difficult to establish a sensitive human-model that can predict operator's action under time-varying circumstance.

Issue I1 appears to come from two points: 1) it is difficult for a machine to measure all types of environmental information that was perceived by human, and 2) the machine recognition of causality of events is hard since many factors and events affect each other. For a general machine operation, as shown around the right area in Fig. 1, three types of statuses, that are the machine status (M-status, say, q_M), the environment status (E-status, q_E), and the task substance (T-status, q_T), can be considered [3]. It is not easy to determine which status is a reason or a result since these statuses influence one another.

Concerning the issue I2, a discrete event modeling is proper to describe human cognitive behavior, and several effective methods such as GOMS [4] and Therblig [5] were developed. It is, however, unexpectedly difficult to embed time factors into a frame of these discrete models. As attempts to solve this issue, many researchers proposed a wide variety of hybrid systems. Suzuki et al. presented the stochastic switched ARX model, and they applied the method to identification of driver's behavior [6]. The model consists of continuous-time linear subsystems, and the switching probabilities of the subsystems are trained using an EM-algorithm. Kawashima et al. proposed a hybrid system involving a finite state automaton and multiple liner dynamical systems [7]. They applied it to a segmentation of multimedia timing structure of human speech using the lip image and voice. Kawashima's approach would be able to be utilized for human intention estimation. Although effectiveness of those methods were demonstrated, the applicable conditions are limited: specifically, the number of the discrete modes were small; the continuity of the discrete state transition was implicitly demanded. Additionally, those methods just classified discrete chunks from the input-output data of the human system, that is, those existing methods did not estimate human internal status that caused the state transition of the discrete modes. The human internal status itself is intention, and it is an essential factor for decision of user's action; however, true estimation of intention is difficult. Omori et al., alternatively, presented an approach for human-robot interaction without intention estimation [8]. There, the self agent (= a machine) actively approaches another agent (= a human) to induce actions by creating a sequence of actions that are easy to be interpreted by the other. Although such studies are interesting, the internal status, intentions, should be able to be estimated to achieve advanced machines to support human operators.

Therefore in the present paper, a new method to estimate operator's intention is proposed without modeling the human thinking process and discrete event framework. Issue I1 is settled by Self-Organizing Map (SOM) which is an effective clustering method for large data. Since the SOM technique can compress multi-dimensional information into two dimensional map by keeping original topological information, the topological information can be utilized as Bayes probability.

This work is supported by a Grant-in-Aid for Scientific Research (A) of the Japanese Ministry of Education, Culture, Sports, Science and Technology.

S. Suzuki is with School of Science and Technology for Future Life, Department of Robotics and Mechatronics, Tokyo Denki University, 2-2 Kanda-Nishiki-cho Chiyoda-ku Tokyo 101-8457, Japan ssuzuki@fr.dendai.ac.jp

F. Harashima is with Tokyo Metropolitan University, 1-1 Minami-Osawa Hachioji-Shi, Tokyo 192-0397, Japan f.harashima@ieee.org

And, transition of intentions is computed probabilistically by the particle filtering technique with the Bayesian probability embedded in the SOM.

The main contribution of the present paper is to propose an algorithm of such intention estimator using Bayes filtering with the SOM technique. The effectiveness of the presented intention estimator was verified by applying to a remote operational task. And second purpose of this paper is to obtain findings to improve the presented intention estimator through the experimental verification.

An organization of this paper is as follows. Section II explains a structure of the intention estimator and preliminary explanation of the Bayes filter. In Section III, a particle filtering algorithm to implement the proposed intention estimator is presented. In Section IV, a remote operation experiment system and its test are mentioned. Section V shows analyses of the applied example and discusses the results. Last Section V is conclusion and discussion.

II. Basic concept of the intention estimator

A human model is assumed in this paper as follows.

$$z_t = f(z_{t-1}, s_t) \tag{1}$$
$$\rho_t = g(z_t), \tag{2}$$

where the subscript $t \in \mathcal{N}$ is a time counter, the function f is a state transition function describing change of the intention z, and the function g corresponds to selection of operational commands based on the intention z. Referring *spotlight models* explaining consciousness in psychology, mathematical expression z is defined as a vector $z \in R^{n_z}$ of which element corresponds to one intention strength of one operation action. According to the *spotlight models*, several types of consciousness exist simultaneously inside human brain, and one of them floats from unconsciousness level as a conscious awareness [9]. An origin of the name comes from circumstance on the stage where only actors (or actress) who is lighted by a spot light can be seen from spectators. Since there are several operational modes in case of a general machine operation, the vector-form expression of intentions can fit a concept of *spotlight models*. Another merit of the vector style is feasibility that enables to analyze dynamically each type of transition.

Considering above-mentioned discussion, probabilistic distribution $bel(z)$, that is a belief of the intention z, will be estimated using a technique of the Bayes filtering. Basic algorithm for estimation of $bel(z)$ is explained below.

Algorithm 1 : Bayes filter

$$\overline{bel}(z_t) = \int p(z_t|s_t, z_{t-1}) \cdot bel(z_{t-1})\ dz_{t-1} \tag{3}$$
$$bel(z_t) = \eta \cdot p(\rho_t|z_t) \cdot \overline{bel}(z_t) \tag{4}$$

This algorithm is defined with iterative equations that are computed from time $t = 1$ to the final time T. $p(z_t|s_t, z_{t-1})$ corresponds to a probabilistic distribution of a transition of intention from z_{t-1} to z_t given input s_t. $p(\rho_t|z_t)$ is the conditional probabilistic distribution of judgment that outputs ρ_t if the intentions z_t happens to be true. Eq.(3) is a *prediction* to obtain a belief $\overline{bel}(z_t)$ at the time of t. Eq.(4) is called *measurement update* and adjusts the prediction $\overline{bel}(z_t)$ by considering a probability $p(\rho_t|z_t)$. Via this update, a new belief $bel(z_t)$ at time t is obtained. η is a so-called Bayes normalization constant. In the proposed approach, mapping relations shown below are acquired approximately with a SOM technique.

$$z_t \leftarrow SOM_f(z_{t-1}, s_t) \tag{5}$$
$$\rho_t \leftarrow SOM_g(z_t) \tag{6}$$

Figure 1 shows a block diagram to explain an operator model and the intention estimator. As shown at the upper left area in the figure, prediction $\overline{bel}(z)$ is computed through SOM_f using input s, the prediction is updated through SOM_g by referring ρ, and $bel(z)$ was obtained. Finally estimated intention $\hat{z}$ is derived.

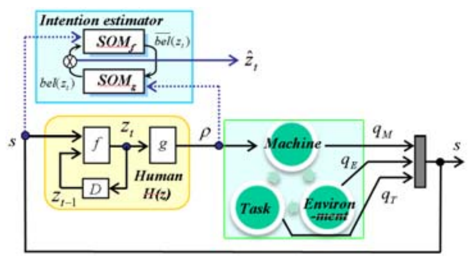

Fig. 1. Block diagram involving a human model and the intention estimator

III. Implementation by particle filter algorithm

Bayesian computation described by Eqs.(3)(4) is implemented using the particle filtering technique [10]. Assuming the number of particles is M, the $m(= 1, \cdots, M)$ th particle of the belief $bel(z_t)$ at time t is described as ${}^{B}z_t^{[m]}$. As a preparation to generate their particles, standard derivations $\sigma_i\ (i = 1, \cdots, n_s)$ of the sequence data $\{s\} \in R^{n_s \times T}$ in each element are computed, where n_s is the size of vector s. Next, SOMs (5) and (6) are trained using normative data of certain identified intentions. Specifically, SOM_f for the state transition function f is trained using the input vector sequence including time-series data $\{s\}$ and $\{z\}$, and then the SOM reference vectors ${}^{f}\xi_i (i = 1, \cdots, L_f)$ are obtained, where L_f is the number of all nodes in SOM_f. Similarly, training the SOM_g for the measurement function g by using sequences $\{\rho\}$ and $\{z\}$, the other SOM reference vectors ${}^{g}\xi_i (i = 1, \cdots, L_g)$ are obtained, where L_g is the number of all nodes in SOM_g. Here, $\{s\}$ and $\{\rho\}$ are made from the experimental logging data, and $\{z\}$ is prepared by an analyst who monitors the record of the expert's operation. The analyst differed from the participant who performed the experimental task. The details will be explained in section IV-B.

Processes of computations in the SOM particle filtering algorithm are explained below by a pseudo-code. The algorithm consists of two phases: phase I for predictive computation and phase II for measurement updating. Below, notation Z_t is used to express a set of particles as $Z_t := \{{}^{B}z_t^{[1]}, {}^{B}z_t^{[2]}, \cdots, {}^{B}z_t^{[M]}\}$.

Algorithm 2 : SOM particle filter

1: $initialize\ Z_1,\ \ prepare\ \{s\}\ and\ \{\rho\}$
2: $for\ t = 2\ to\ T\ do$
3: $\quad Z_t = \emptyset$

phase I
4: $\quad for\ m = 1\ to\ M\ do$
5: $\qquad extract\ {}^{B}z_{t-1}^{[m]}\ from\ Z_{t-1}$
6: $\qquad generate\ {}^{B}\bar{z}_t^{[m]} \sim p(z_t|s_t, {}^{B}z_{t-1}^{[m]})\ by\ SOM_f$
7: $\qquad w_t^{[m]} = p(\rho_t|{}^{B}\bar{z}_t^{[m]})\ using\ SOM_g$
8: $\quad endfor\ m$

phase II
9: $\quad for\ i = 1\ to\ M\ do$
10: $\qquad draw\ m'\ with\ probability \propto\ w_t^{[m]}$
11: $\qquad add\ {}^{B}\bar{z}_t^{[m']}\ to\ Z_t$
12: $\quad endfor\ i$

13: $endfor\ t$

At the phase I, a best-matching-node (BMN) that is close to a combination of the input s_t and the particle ${}^{B}z_{t-1}^{[m]}$ is searched from SOM_f, and the predicted particle ${}^{B}\bar{z}_t^{[m]}$ at time t is picked up by the BMN index (line 6). At this step, the predicted belief is expressed by a set consisting of M nodes generated by adding perturbation into input s_t, and this perturbation technique is similar to normal particle filtering approach. Measurement probability $w_t^{[m]}$ corresponding to these predicted particles is computed using SOM_g for next re-sampling process on after-mentioned phase II (line 7). At the phase II, a particle number m' is chosen in proportion to the measurement probability $w_t^{[m]}$ of each particle (line 10), and the predicted ${}^{B}\bar{z}_t^{[m']}$ which is indicated by the chosen number m' is picked up into a new set Z_t as a next time particles. Then, new M particles are re-sampled according to the measurement probability $w_t^{[m]}$. Repeating phase I and phase II till the final time T, $\{bel(z_t)\}$, that is a time-series $\{Z_t\}$ of the belief of the intention z, are obtained. Below, main parts of the details of Algorithm 2 are explained.

Line 6: Prediction

L6-1: Preparation of perturbation input

Random sample point ${}^{\Delta}s_t$ that obeys a standard deviation σ_i around s_t is computed for all $m(= 1, \cdots, M)$. Specifically, using a random value Δ, the sample point is computed as

$$\begin{aligned} {}^{\Delta}s_t^{[m]} &= s_t + \Delta \\ &\Delta := rand([-1,1], \sigma), \end{aligned} \tag{7}$$

where $rand([-1,1], \sigma)$ is a function that yields pseudo-random value in the range of $[-1,1]$ under standard deviation σ [11].

L6-2: Search of most likelihood node

Using reference vectors ${}^{f}\xi$ of the SOM_f, the BMN of a particle m, that is ${}^{f}c_t^{[m]} \in \{1, \cdots, L_f\}$, is found as

$${}^{f}c_t^{[m]} = arg\min_i \left\{ \left\| \begin{bmatrix} {}^{\Delta}s_t^{[m]} \\ {}^{B}z_{t-1}^{[m]} \end{bmatrix} - {}^{f}\xi_i \right\| \right\}.$$

L6-3: Extraction of prediction state

A candidate of an intention involved in the particle that corresponds to prediction $\overline{bel}(z_t)$ is extracted from a reference vector of the BMN predicted at the L6-2 step:

$${}^{B}\bar{z}_t^{[m]} \Leftarrow {}^{f}\xi_i(n_s + 1 : n_s + n_z),\ i = {}^{f}c_t^{[m]},$$

where an operation described by a parenthesis in the above RHS indicates an extraction of elements of the vector components.

Line 7 : Computation of measurement probability

Finding a BMN of ${}^{B}\bar{z}_t^{[m]}$ by using SOM_g, reference vectors that belong to certain region around the BMN on the SOM_g plane map are investigated. The number of nodes whose reference vectors correspond to the measured command ρ_t appears to be proportional to the post-measurement probability; hence, the measurement probability is computed from the number of such nodes. For the re-sampling process at after-mentioned phase II, an information of such nodes is registered into a roulette array W. Numbering of particles is recorded in the array, and the number of the numbering is determined in proportion to the amount of the corresponding nodes. Followings are the details.

L7-1: Initialization

Reset the roulette array as $W = \emptyset$.

L7-2: Search of most likely node

Using the reference vectors ${}^{g}\xi$ of SOM_g, a node ${}^{g}c_t^{[m]}$ that is most close to the measured ρ_t and predicted status ${}^{B}\bar{z}_t^{[m]}$ is found by

$${}^{g}c_t^{[m]} = arg\min_i \left\{ \left\| \begin{bmatrix} \rho_t \\ {}^{B}\bar{z}_t^{[m]} \end{bmatrix} - {}^{g}\xi_i \right\| \right\}.$$

L7-3: Investigation of region around the most likely node

Computing a coordinate value (u_c, v_c) of the node ${}^{g}c_t^{[m]}$ on the SOM_g plain map, reference vectors ${}^{g}\xi_{i'}$ of nodes that locate inside a square-like region $R_{{}^{g}c_t^{[m]}}$ are investigated, where the length of side and the center of $R_{{}^{g}c_t^{[m]}}$ are $(2L_r + 1)$ and (u_c, v_c), respectively. Extracting from $\xi_{i'}$ an element that corresponds to operation command, its element is described as $\hat{\rho}^{(i')}$, that is

$$\hat{\rho}^{(i')} \Leftarrow {}^{g}\xi_{i'}(1 : n_\rho),\ i' \in R_{{}^{g}c_t^{[m]}}.$$

L7-4: Registration to the roulette array

The number of the variables that belong to $\hat{\rho}^{(i')}$ ($\in$

$\mathcal{R}, \sum i' \leq (2L_r+1)^2)$ and are close to $\rho_t(\in \mathcal{I})$ is counted as follows, and the number is defined as l.

$$l = \left\{ \sum i' \mid round(|\hat{\rho}^{(i')} - \rho_t|) = 0 \right\}$$

Next, 'the number, m' is registered l times into the array W additionally.

$$W \Leftarrow W + \{m\}^l$$

Line 10: Selection

Since the numbering of particle that holds higher measurement probability has been registered in W more times, such particles are re-sampled again with high rate by random selection. Hence, generating a random integer r within a range of $\{1, \cdots, n_W\}$, a number m' that was registered on the r-th element of the array W is drawn, where n_W is a length of W.

$$\begin{aligned} m' &= W(r) \\ & \quad r = rand(\{1, \cdots, n_W\}). \end{aligned} \tag{8}$$

Line 11: Reentry

m'-th particle is re-registered as one of new particles for next step as

$${}^B z_t^{[m]} \Leftarrow {}^B \bar{z}_t^{[m']}.$$

Algorithm 2 yields M particles ${}^B z_t^{[m]}$ every iteration time. Since the belief is expressed by a distribution of those M particles, an estimation of intention, say $\hat{z}$, is defined by averaging these M particles as a representative variable to show the time transition of estimated belief:

$$\hat{z}_t = \frac{1}{M} \sum_{m=1}^{M} {}^B z_t^{[m]}.$$

Conceptual diagram of the SOM-Bayes filtering is shown in Fig. 2.

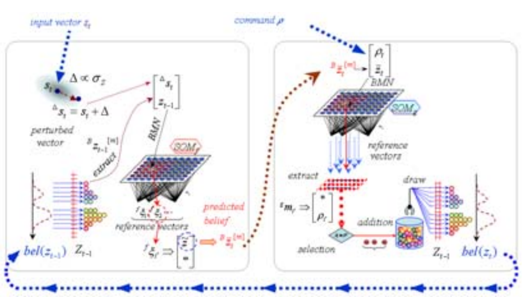

Fig. 2. Diagram of the SOM-Bayes filtering

IV. Remote operation of radio controlled construction equipments

A. Experimental set up

An experimental system with radio-controlled model construction machines [12] was devised as a human-operation test for verification of the proposed method. The task is a basic soil excavation. Wireless cameras on the excavator and the truck captured video images, and displayed on monitors for the operator, as shown in Fig. 3 (a). The operator manipulated both the excavator and the truck at one's own discretion. The operator read the instruction manual for the usage of the controllers and for the purpose of task before the examination, and started the operation without watching any other person's operation. Figure 3 (b) shows the overview of the work area. The field consists of the motorable road, the restricted area, three drilling sites, and one unloading site. The excavator and truck were put at their starting position at the beginning of trial. The operator moved the machines to the drilling site, collected the sample pieces with the excavator, loaded the pieces on the truck bed, and carried them to the unloading site by the truck. Written consent and ethical approval of the participant (aged 21) were obtained before the examinations. Three trials were repeated for three days, hence the total were 9 trials.

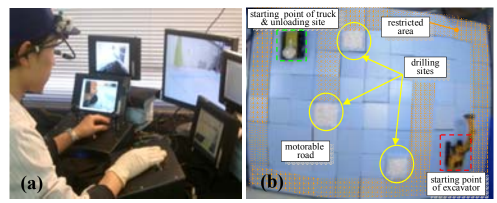

Fig. 3. Remote operation console (a), and overview of the work area (b).

B. Preparation for the SOM-Bayes intention estimator

Crawler velocities for the excavator and the truck were controlled by two sliders with hands. The velocity commands were converted into the crawler operation mode κ_c as $\kappa_c = \{0\ (stop),\ 1\ (forward),\ 2\ (left\ f.\ steer),\ 3\ (left\ pinwheel),\ 4\ (right\ b.\ steer),\ 5\ (backward),\ 6\ (left\ b.\ steer),\ 7\ (right\ pinwheel),\ 8\ (right\ f.\ steer)\ \}$, where "*f.*" and "*b.*" are abbreviations of "forward" and "backward", respectively. The bucket arm is manipulated by two cross levers, and the operator commands consisted of the superstructure rotation mode; ${}^e\kappa_r := \{0 : stop,\ 1 : left\ rotation,\ 2 : right\ rot.\}$, the arm mode; ${}^e\kappa_a := \{0 : stop,\ 1 : arm\ bend,\ 2 : arm\ stretch\}$, and the bucket mode; ${}^e\kappa_b := \{0 : stop,\ 1 : boom\ up,\ 2 : boom\ down\}$, respectively. All operation modes, that consist of the truck crawler modes, the excavator crawler modes and the excavator shovel operation mode, were summarized into one variable $\rho \in \{1, \cdots, 42\}$ [12] as

$$\rho := 1 + {}^t\kappa_c + ({}^e\kappa_c \wedge 1)({}^e\kappa_c + 8) + ({}^e\kappa_h \wedge 1)({}^e\kappa_h + 16). \tag{9}$$

Vectors of the machine and environmental status (*i.e.*, q_M and q_E) were chosen by considering position, posture, and geographical relation of the drilling sites and the equipments. The task status q_T was defined as $q_T := \{0 : no\ payload,\ 1 : payload\ on\ bucket,\ 2 : payload\ on\ truck\ bed\ \}$ using the payload status. Refer [13] for details about selection of these status. Time series sequence $\{s\}$ were obtained by combining q_M, q_E and q_T, and the size of vector s became as $n_s = 28$.

Sequence of reference intentions $\{z'\}$ was made through video analysis by a participant who watched the recorded video of the expert's operation. Types of the remote operations for construction equipments and the definitions of elements $(z_1, \cdots, z_{15})$ in intention vectors z are summarized in Table I. The operation modes are classified into three groups: approaching $(z_1 \sim z_6)$, positioning $(z_7 \sim z_{12})$, and special operations (digging, loading, and transporting; $z_{13} \sim z_{15}$). To get rid of any preconceived ideas, this intention analysis was asked a participant who did not know the remote-operation experiment at all. The analyst discerned the operator's intention, as classified in Table I, by watching the video recorded from the ceiling camera, and tried to find the timings of each operation, then sequence $\{z\}$ was made by putting 1 to the corresponding element z_i at the time of the found timing.

TABLE I
CLASSIFIED OPERATIONAL INTENTIONS, THEIR VARIABLES, AND LABELS

variable	label	meanings
$z_1 \sim z_3$	T/A-*	truck approach to the *-drilling site
$z_4 \sim z_6$	E/A-*	excavator approach to the *-drilling site
$z_7 \sim z_9$	T/P-*	truck positioning around the *-drilling site
$z_{10} \sim z_{12}$	E/P-*	excavator positioning around the *-drilling site
z_{13}	E/D	excavator digging
z_{14}	E/L	excavator loading of the payload
z_{15}	T/T	truck transport with the payload

(*=a,b,c : identifier of the drilling site)

Computation of the SOM training was performed using SOM_PAK [14]. Each component of the input vectors was normalized into the $[-1, 1]$ range by using the maximum and minimum values of the time series data. Since it is preferable for the horizontal and vertical sizes of the rectangular map of SOM to be chosen in proportion to the ratio of two square roots of the first and second maximum eigen-values of the covariance matrix of input vectors [15], the sizes of the SOM lattice (u_{dim}, v_{dim}) were decided as $(50, 25)$ and $(80, 20)$ for SOM_f and SOM_g, respectively. Hence, the number of nodes L_f and L_g are 1250 and 1600, respectively. A bubble type was chosen for a neighborhood kernel function. On the learning process, a fine tuned computation was performed after the rough tuned one was computed. The learning rate & learning length were specified as 0.05 & 2000 and 0.02 & 1.5 million, respectively, so as to meet such requirement that the learning length is more than 500 times of the number of nodes [16].

Initial state of particles were specified as ${}^{B}z_1^{[m]} = [\ 0\ 0\ 0\ 1\ 0\ \cdots\ 0\ ]^T$ for all m since it was obvious that an initial operation in the experiment was E/A-a (that means 'excavator approach to the a-drilling site'). The number of particles were specified as $M = 1000$.

V. Analyses

Concerning the intention discerned by a human analyst, say $\tilde{z}$, and the other intention estimated by the proposed SOM-Bayes method, $\hat{z}$, the transitions are shown in Fig. 4. Each graph shows the change of each intention mode from T/A-a (z_1) to T/T (z_{15}). $\tilde{z}$ and $\hat{z}$ are drawn by the blue and red lines, respectively.

A. Qualitative analysis

Investigating the overlap, timing, strength, and types of the intention qualitatively by comparing $\hat{z}$ against $\tilde{z}$, the following tendencies were found.

1) Concerning *approaching*, both case of the truck and the excavator (T/A-*, E/A-*) were identified correctly.
2) Concerning *truck transport* (T/T), starting timing of the estimation was delayed.
3) Periods identified as *truck positioning* (T/P) were longer.
4) Periods identified as *excavator loading* (E/L) were also long, and their level were strong.

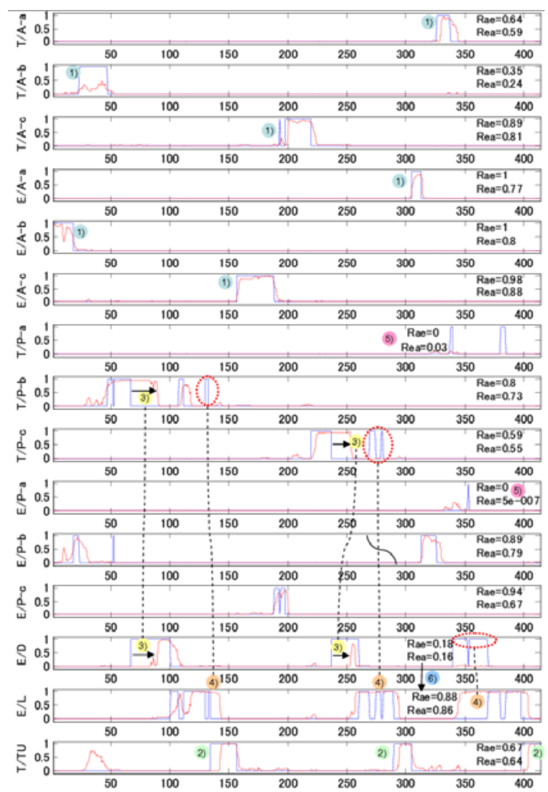

Fig. 4. Transitions of estimated intentions (red) and discerned ones (blue)

For easier understanding correspondence between graphs and above-mentioned results, numbering of items in the results were written in Fig. 4. Reason for the success of result 1 appears to come from large change in machine's state. The reason of delay indicated by result 2 is that analyst discerned early the initiation of T/T, in short, the analyst regarded the end of E/L action (that occurred before the T/T action) as a start of T/T. Result 3 was found by investigating which status changed at the same timing of E/D in $\hat{z}$ (that were about 85[s], 255[s]). Specifically, timing detected by the estimator was synchronized with vertical change of the bucket while analyst's $\tilde{z}$ was changing from the bucket operation in a front-back direction before the vertical manipulation. This insights give us thinking that judgement of human analyst might be inadequate and that the estimator by the machine algorithm appears to identify operator's intention more objectively. From result 4, the estimator seems to indicate that truck's positioning during the payload loading operation includes the E/L operation. Such interpretation of 'T/P is a part of E/L' can be acceptable. From a different angle, this result highlights difficulty such that criterion to classify intentions is ambiguous.

B. Quantitative analysis

Computing the matching ratios using $\tilde{z}$ and $\hat{z}$ by the following indexes, the timing and strength of the estimated intentions were investigated.

$$R_{a|e} \quad := \quad \sum^{all\ t'} \tilde{z} / \sum^{all\ t} \tilde{z}, \quad t' = \{t \mid \hat{z}_t > 0.3\ \} \quad (10)$$

$$R_{e|a} \quad := \quad \sum^{all\ t'} \hat{z} / \sum^{all\ t'} \tilde{z}, \quad t' = \{t \mid \tilde{z}_t = 1\ \} \quad (11)$$

The constant 0.3 in Eq.(10) is a threshold parameter. $R_{a|e}$ is a ratio of time that the analyst found same type of intention against a time period of the other intention identified by the estimator. $R_{e|a}$ shows how strongly the estimator can identify same type of intention during the period recognized by the analyst. Values of these indexes are written at the right side in each graph shown in Fig. 4. Although a weak tendency of $R_{a|b} > R_{b|a}$ can be found, there is not discriminative difference depending on types of intentions between these two indexes. Hence, difference of two indexes were not cared but the differences from the types of intentions were checked. Then, the following results was found.

5) Estimated $\hat{z}$ for T/P-a and E/P-a were not detected.
6) The matching ratios of E/D $(R_{a|b} = 0.176, R_{b|a} = 0.16)$ was small compared to the others.

Reason of result 5 might come from a geographical dependency of the drilling site A. Actually, the truck passed through the road around the site A six times while the excavator reached there once; hence, a probability of the excavator's emergence was low and the excavator's operation might be difficult to be classified. Since result 6 arises from festinate discretions of the analyst, as explained in result 3, the estimation was not necessarily inaccurate. On the other hand, since the matching ratios for E/L are as high as 0.88 and 0.86, respectively, an existence of an inclusion relation such as $E/D < E/L$ are shown. These facts also support the qualitative result 4.

C. Discussion

From afore-mentioned analyses, drawbacks of the proposed algorithm and the estimated results are summarized as the following findings.

F1) Classification of intentions based on individual sense and experience on human analysis is ambiguous. We should think that such ambiguousness cannot be removed clearly on the classification work. It is required to verify believability of the normative data of intentions.

F2) Estimation of intention is strongly affected with factors for foreseeing; a scenario of whole task and causality of events should be considered to improve the proposed algorithm.

F3) At the stage of the SOM training, the input data should be modified to enhance a fixation of the related network in the SOM structure for significant but rare events.

Finding F1 is a corollary from result 4. It was indicated that types of intentions shown in Table I were not perfectly adequate classification and that their types contained certain inclusive relation. The present authors rediscovered that it was difficult to make normative data for an intentions estimation. F2 indicates a drawback of the proposed method, and is obtained from results 2 and 3. That is, a prediction by Eq.(7) utilizing stochastic perturbation was insufficient. Problem shown in F3 was deduced from result 5. The proposed algorithm is a so-called frequentism method consisting of Bayes estimation; hence, events with low frequency are difficult to be classified. At the learning phase for SOM, an adjustment of input data for learning is required according to significance level of each event.

It might be thought that the presented SOM-Bayes estimator may identify not operational intention but operational sequence; however, it is misunderstanding. The SOM-Bayes identifies stochastically the inner status which determines the output commands, and such inner status is called intentions. We considered a simultaneous existence of intention modes by using multiple variables. These points are the biggest difference between the SOM-Bayes estimator and existing estimators of operational sequence.

VI. Conclusion

A new method to estimate operator's intention was presented by utilizing the SOM clustering technique and the Bayes estimation. Characteristic of this method is to form a state transition relation of intentions by using SOM. In the presented method, troublesome preparation to specify informations in human cognitive activity on machine operations is not necessary. Applying the presented method to the remote operation task, the effectiveness was verified through qualitative and quantitative analyses, and several issues and benefits were confirmed. As a result, the following findings were obtained: 1) objective attention should be paid to the classification of reference intentions for the SOM training, 2) it is desirable to take a scenario of events and the causality into consideration to predict state transition on the Bayes filtering computation, and 3) an input data for the SOM training should be modified according to meanings of events and actions. Consideration and improvement concerning these issues are the remained work in the future.

Acknowledgment

The experiment was supported by many participants who embraced our requests kindly. The present author appreciates their cooperation.

References

[1] C. M. Bishop, “Pattern Recognition and Machine Learning,” Springer, 2006.

[2] F. Harashima and S. Suzuki, “Intelligent Mechatronics and Robotics,” keynote speech, in *Proc. the 2008 IEEE International Conference on Emerging Technologies and Factory Automation (ETFA2008)*, Hamburg, Germany, Sept. 2008, in CD-ROM.

[3] S. Suzuki and F. Harashima, “Segmentation and analysis of console operation using self-organizing map with cluster growing method,” in Proc. of The 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2009), October 11-15, St. Louis, USA, pp.4875–4880, 2009.

[4] W.D. Gray, B.E. John, and M.E. Atwood, “Project Ernestine: Validating a GOMS Analysis for Predicting and Explaining Real-World Tasks Performance,” *Human-Computer Interaction*, vol. 8, pp 237-309, 1993.

[5] B. Price, “Frank and Lillian Gilbreth and the Motion Study Controversy, 1907-1930,” A Mental Revolution: Scientific Management since Taylor, Daniel Nelson, ed. The Ohio State University Press. 1990.

[6] T. Suzuki, S. Sekizawa, *et al.*, “Modeling and Recognition of Human Driving Behavior based on Stochastic Switched ARX model,” in *Proc. of the 44th IEEE Conference on Decision and Control, and the European Control Conference 2005 (CDC/ECC2005)*, Seville, Spain, pp. 5095-5100, 2005.

[7] H. Kawashima and T. Matsuyama, “Multiphase Learning for an Interval-Based Hybrid Dynamical System,” IEICE Trans. Fundamentals, vol.E88-A, no.11, pp. 3022–3035, 2005.

[8] T. Omori, A. Yokoyama, H. Okada, S. Ishikawa, and Y. Nagata, “Computational Modeling of Human-Robot Interaction Based on Active Intention Estimation,” Lecture Notes in Computer Science: Neural Information Processing, Vol. 4985, pp. 185–192, 2008.

[9] B. J. Baars, *A cognitive theory of consciousness*, Cambridge University Press, New York, 1988.

[10] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, MIT Press, Cambridge, MA, 2005.

[11] G. Winkler, “Image Analysis, Random Fields, and Dynamic Monte Carlo Methods,” Berlin, Springer Verlag, 1995.

[12] S. Suzuki, and F. Harashima, “Skill analysis focused on the hand discrete movement for machine manipulation,” in *Proc. of the 13th IEEE Int. Conf. of Emerging Technologies and Factory Automation (ETFA2008)*, Hamburg, Germany, pp 156–163, 2008.

[13] S. Suzuki, “Visualization of task switching strategy of machine operation,” *in the 2009 IEEE International Conference on Networking, Sensing and Control (ICNSC2009)*, Okayama City, Japan, pp.156–163, 2009.

[14] http://www.cis.hut.fi/research/som_lvq_pak.shtml

[15] T. Kohonen, Self-Organizing Maps, *Springer-Verlag*, Heidelberg Berlin, 1995.

[16] T. Kohonen, “The self-organizing map,” *Proc. of the IEEE*, vol. 78, no. 9, pp 1464-1480, 1990.