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

System of Recognizing Human Action by Mining in Time-Series Motion Logs and Applications

Yihsin Ho, Kota Nakamura, Tomomi Shibano, Eri sato-Shimokawara and Toru Yamaguchi

Abstract—**For solving the "Digital divide" problem and developing a system which understands human, the data mining technology is being developed. Moreover, within the concept of Kukanchi, the system can not only collect information, but also provide service to user. In this paper, the authors show two experiments which constructed based on "Kukanchi". These Kukanchi systems include both techniques of action recognition and mining in time-series motion logs. In the first experiment, the authors constructed a service robot system, which makes networked robots exchange and provide the information to customers by recognizing their actions, to show the basis concept of the paper. In the second experiment, the authors show the possibility of using interactive data mining, which uses visualization techniques to help human add judgments, to improve the results of the first experiment. The two experiments show the utility and possibility of Kukanchi and mining in time-series motion logs which supports the system to understand human and help people directly.**

Index Terms—Human action, mining, motion logs

I. INTRODUCTION

Nowadays many robot systems were developed and widely used in our world. The progress of technology makes our life more convenient than ever. But the benefit is only helping for those people who know and are familiar with using application systems. For those people who are unfamiliar with system usages, the application systems are still hard to understand and difficult to use. We call this problem "Digital divide" and the problem also shows the importance of developing a system that can understand human.

For developing the system's capability of understanding human, the data mining technology is also developed. Data mining can discover the knowledge by the data which is collected by human or other equipments. If the information of users can be collected and processed by data mining, the discovered knowledge can support the system to understand human. For example, the research results of mining on web and POS (Point of sales) have been utilized in marketing and used to provide services. But, for using those mining techniques, there always is a person to input the necessary data and operate the equipments.

In this paper, the system obtains the characteristic behaviors of human by processing the images which are captured by stereo cameras. By mining in time-series motion logs and extraction of association rule, the authors acquire the knowledge to recognize the actions. And the results can be used as a human action forecast data to help the Kukanchi system distinguishes user's situations and intentions, and then provides service appropriately.

II. NETWORKED ROBOT AND KUKANCHI

In the speech of the first conference of ISRR, Professor Winston of MIT defined the robotics as follow: "Robotics is the intelligent connection of perception to action." In other words, as the definition, the behaviors of the intelligent robot come from the perception of the robot. This can be thought of the exact definition of intelligent robots. Recently, the word "robot" has been used for broader meaning than it used to be. Robot also means the intelligent system which applies the robot technology to support real world. ATR-centered group [1] that researches the networked robot technology separated the robot into 3 types. The ideal is very similar to the authors' image of networked robot, but the concept maybe different [2]. The first type is virtual type robot which is included into Smartphone or etc., the second type is the unconscious type robot that focuses on information collection and situation recognition, and the third type is the visual type robot which is for handling the situation and directly communicating with human. The three type of robots work cooperation and exchange the information to each other. By sharing information with each other or the environment, the three type robots can cooperate and provide service. This way of thinking is called "Kukanchi". The information which is collected from the Kukanchi system can also be thought of a useful data for utility, for example, the information can use in data mining to discover the valuable knowledge.

III. MINING TECHNOLOGY OF COMBINING HUMAN AND COMPUTER

The mining techniques which be developed recently include not only data processing, but also a series of techniques like data preparation, data cleansing, visualization techniques and decision making. Every step of the process is important. Especially visualization techniques that help humans make their judgments. In the mining processing, humans add their judgments at the visualization step for getting the most suitable data or data cleansing function for processing mining. After visualization step, the next step is processed for finding the valuable knowledge. If the data or function is not help for discovering knowledge, the visualization step will be processed again, until useful knowledge is found. Moreover, there are some important

things of discovering knowledge by data mining, like the appropriate judgments of human, the performance of processing interactive data mining and the step of keeping mining process is implemented until the appropriate knowledge is found.

For exactly explaining the concept of Kukanchi and the techniques of recognizing human action by mining, the authors show two experiments and the mining technique in chapter 4 and 5, and make the conclusion in chapter 6.

IV. THE FIRST EXPERIMENT: THE ROBOT SYSTEM OF A CLOTHES STORE AND EXTRACTION OF ASSOCIATION RULE

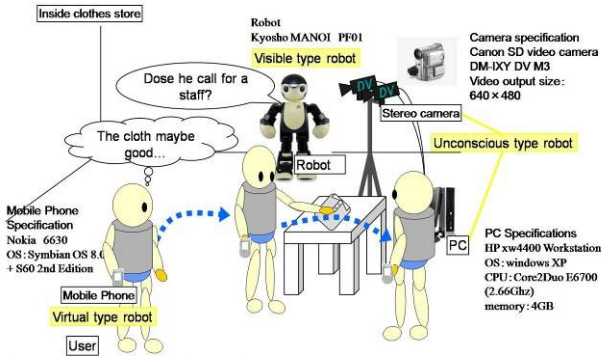


Fig.1. Composition of experiment system.

In this chapter, the authors intend to develop a system using in a clothes store as an experiment. Fig.1 shows the composition of system. The system recognizes human actions by mining in time-series motion logs from images which captured by stereo cameras, and provide the information to customers. The authors focus on the human motions of customers to construct the system. The system captures human motions by processing the images from stereo cameras. The human motions include taking, walking, stop and turning in front of the merchandise. The authors take the four human motions as the characteristic behaviors. The large data of human motions can be classed by the four characteristic behaviors and the class can also make the data become simple, especially in processing data cleansing.

A. Action recognition

For recognizing human actions, the authors use the motion capture equipment, which is named Vicon, to get images first. Then, the authors use V-Space, which is developed by Yamaguchi Lab., as the image processing software to capture motion data. V-Space uses the RGB(R: Red, G: Green, B: Blue) color space as the basic color space to process the images and can trace the object by its color. The system analyzes the colors in the original images which are captured by the camera, and calculates the weight points and the coordinates by the color pixel, and then, saves the point and coordinate data into database. By repeating the process of calculating the weight points and coordinates after a period of time, the system can use the data to trace the object, and the user can find the trajectory which he needs by the color data of the target. Moreover, the system also includes background subtraction technique to improve the capability of detecting moving objects or the object which is with complex background.

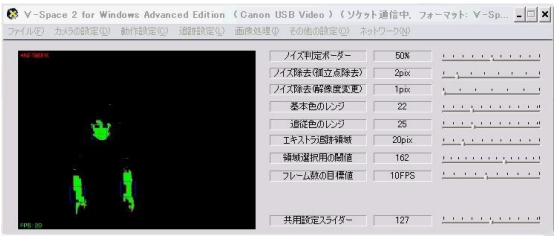


Fig.2. The experiments view of V-Space.

After capturing the human motion data from V-Space, the analysis of human actions can be implemented. The rules of characteristic behavior recognition are as follow.

1) Turning : If the variable of angle between parallel plane and the ground coordination of the head is over the threshold value every 0.1 second, it can be judged as “turning”.

2) Taking : If the coordination of right or left hand is at the position of merchandises every 0.1 second, it can be judged as “taking”.

3) Walking and Stop : If the variable of vector-valued between parallel plane and the ground coordination of the head is over the threshold value every 0.1 second, it can be judged as “walking”. Otherwise, it can be judged as “stop”.

Because the data may be captured by other equipments which like stereo cameras, the accuracy of data which is obtained from stereo cameras is less than Vicon. The results of the rules, which are described before, may include noise which affects the behavior patterns that is created and used for recognizing human actions. So, the authors implement the visualization techniques on the human motions which are obtained from stereo cameras to reconstruct the rules of characteristic behavior recognition. The new rules are shown in Fig. 3

Here the authors explain the condition of don’t care (the authors use the mark * to represent). Before creating the result by the rule that are illustrated in Fig. 3, if the variable of vector-valued for judging the walking and stop states is under the threshold value, then the data will be processed by trigger. And the function, which is for judging “stop” and “walking”, will work on the four vectors that are created after the process of trigger. After recognizing the characteristic behaviors by the rules, although the system can obtain a more stable motion from walking state, the two problems that the accuracy problem of stereo cameras and the noises of the head shaking in stop state still arise the problem of false recognition. So, * which is the middle of walking and stop, or in other words, ambiguous operation which is like a noise, is applied to identify and separate the conditions, like “walking but nearly stop” and “exactly stop”.

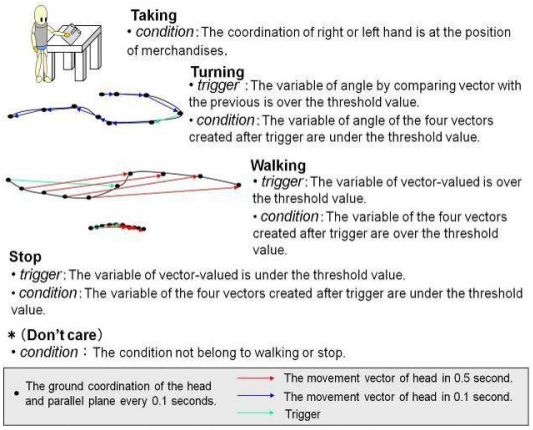


Fig.3. The new rules of characteristic behavior recognition.

B. *Extraction of association rule*

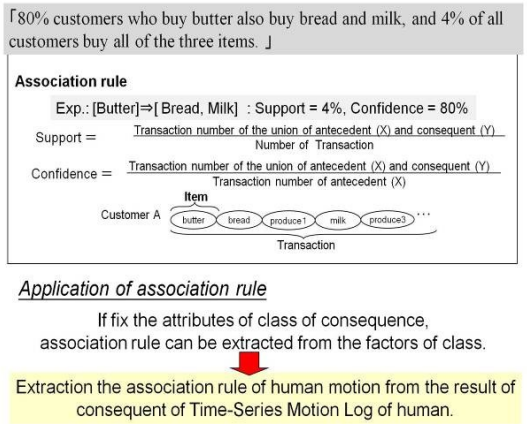


Fig.4. The Example and formula of extraction of association rule.

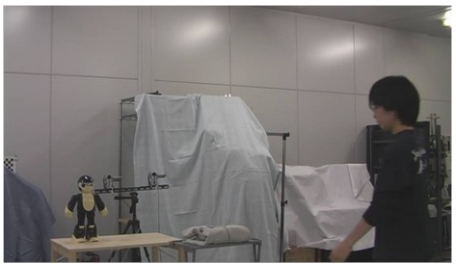


Fig.5. The arrangement of the devices of the experiment.

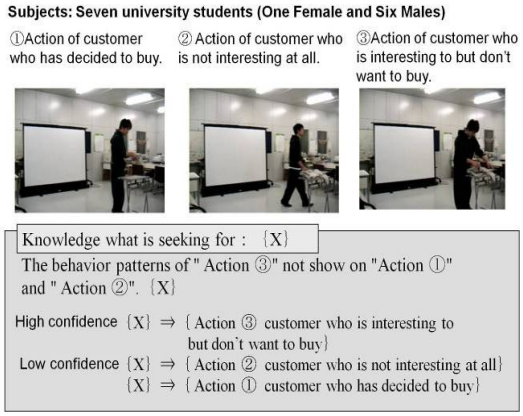


Fig.6. The behavior patterns and knowledge discovery.

Generally, association rule is known as marketing basket which is shown in Fig.4. But if we fix the attributes of class of consequence, association rule still can be extracted from the factors of class [3]. In this paper, the target is to extract the association rule of human action from the consequence of human behavior patterns, and apply the result into system for recognizing human actions and offering services at the appropriate time.

Fig.7. The process of classifying and extraction of association rule.

For preceding the experiment, the authors invite seven university students, one female and six male, as the subjects to recode their motions, and arrangement of the devices in the experiment as illustrated in Fig. 1 and Fig. 5. The authors capture the images of the subjects by Vicon and use the function to judge and analyze the motions of subjects for collecting the data of characteristic behaviors, and then, use the data to create behavior patterns and recognize action. Fig. 6 shows the process of collecting characteristic behaviors and using the behavior patterns to process the extraction of association rule and knowledge discovery.

Because extracting from all the items of time-series motion logs will cost large complexity. Before analyzing the data of behavior patterns, the characteristic behaviors will be classified by those feature actions, as described in Fig. 7.

The process of analyzing the behavior patterns will be easier by the classified data. After the process, the system can calculate the confidence and support by the formula of extraction of association rules to get the result. The Table I, Table II and Table III show the experiment results after classified by those feature actions and the statistics after judging and analyzing the data of behavior patterns.

In this research, the authors focused on the characteristic behaviors of "taking" as the main factor of behavior patterns, and using the behavior patterns to extract association rule. The result {stop, taking}×3, which is shown in Table I, and the sum of taking action＝19, which is shown in Table III, present the discovered knowledge of (3)(Action of customer who is interesting to but don't want to buy), which is shown in Fig.6.

TABLE I
THE DATA OF BEHAVIOR PATTERN: {STOP, TAKING}

	Action of customer who is interesting to but don't want to buy (%)	Action of customer who is interesting to but don't want to buy (%)	Action of customer who has decided to buy (%)	Action of customer who is not interesting at all (%)
	Confidence	Support	Confidence	Confidence
{Stop-Taking}×1	78	33	22	0
{Stop-Taking}×2	78	33	0	0
{Stop-Taking}×3	100	33	0	0
{Stop-Taking}×4	100	33	0	0
{Stop-Taking}×5	86	28	0	0
{Stop-Taking}×6	86	28	0	0
{Stop-Taking}×7	71	23	0	0
{Stop-Taking}×8	57	19	0	0
{Stop-Taking}×9	57	19	0	0
{Stop-Taking}×10	43	14	0	0
{Stop-Taking}×11	29	9	0	0
{Stop-Taking}×12	29	9	0	0
{Stop-Taking}×13	29	9	0	0
{Stop-Taking}×14	29	9	0	0
{Stop-Taking}×15	29	9	0	0
{Stop-Taking}×16	14	4	0	0
{Stop-Taking}×17	0	0	0	0

TABLE II
THE DATA OF BEHAVIOR PATTERN: {STOP, *}

	Action of customer who is interesting to but don't want to buy (%)	Action of customer who is interesting to but don't want to buy (%)	Action of customer who has decided to buy (%)	Action of customer who is not interesting at all (%)
	Confidence	Support	Confidence	Confidence
{Stop-*}×1	54	33	38	8
{Stop-*}×2	70	33	30	0
{Stop-*}×3	67	28	33	0
{Stop-*}×4	63	23	37	0
{Stop-*}×5	63	23	37	0
{Stop-*}×6	67	19	33	0
{Stop-*}×7	80	19	20	0
{Stop-*}×8	75	14	25	0
{Stop-*}×9	75	14	25	0
{Stop-*}×10	100	14	0	0
{Stop-*}×11	100	9	0	0
{Stop-*}×12	100	4	0	0
{Stop-*}×13	100	4	0	0
{Stop-*}×14	100	4	0	0
{Stop-*}×15	100	4	0	0
{Stop-*}×16	100	4	0	0
{Stop-*}×17	100	4	0	0
{Stop-*}×18	100	4	0	0
{Stop-*}×19	100	4	0	0
{Stop-*}×20	100	4	0	0
{Stop-*}×21	100	4	0	0
{Stop-*}×22	0	0	0	0
{Stop-*}×23	0	0	0	0

TABLE III
THE DATA OF BEHAVIOR PATTERN: {SUM OF TAKING}

	Action of customer who is interesting to but don't want to buy (%)	Action of customer who is interesting to but don't want to buy (%)	Action of customer who has decided to buy (%)	Action of customer who is not interesting at all (%)
	Confidence	Support	Confidence	Confidence
Sum of {Taking} = 1	50	33	43	7
Sum of {Taking} = 2	58	33	33	9
Sum of {Taking} = 3	64	33	27	9
Sum of {Taking} = 4	70	33	30	0
Sum of {Taking} = 5	70	33	30	0
Sum of {Taking} = 6	78	33	22	0
Sum of {Taking} = 7	88	33	12	0
Sum of {Taking} = 8	88	33	12	0
Sum of {Taking} = 9	88	33	12	0
Sum of {Taking} = 10	88	33	12	0
Sum of {Taking} = 11	88	33	12	0
Sum of {Taking} = 12	88	33	12	0
Sum of {Taking} = 13	86	28	14	0
Sum of {Taking} = 14	86	28	14	0
Sum of {Taking} = 15	86	28	14	0
Sum of {Taking} = 16	83	23	17	0
Sum of {Taking} = 17	83	23	17	0
Sum of {Taking} = 18	83	23	17	0
Sum of {Taking} = 19	100	23	0	0
Sum of {Taking} = 20	100	23	0	0
Sum of {Taking} = 21	100	23	0	0
Sum of {Taking} = 22	100	23	0	0
Sum of {Taking} = 23	100	23	0	0
Sum of {Taking} = 24	100	23	0	0

The tables also show two results of knowledge discovery that (3) (Action of customer who is interesting to but don't want to buy) is 100% confidence, and the confidence of (1) (Action of customer who has decided to buy) and (2) (Action of customer who is not interesting at all) are both 0%. The result, which is shown in Fig.8, presents the knowledge which be found. And it also proposes the proof that knowledge discovery technique is a useful method.

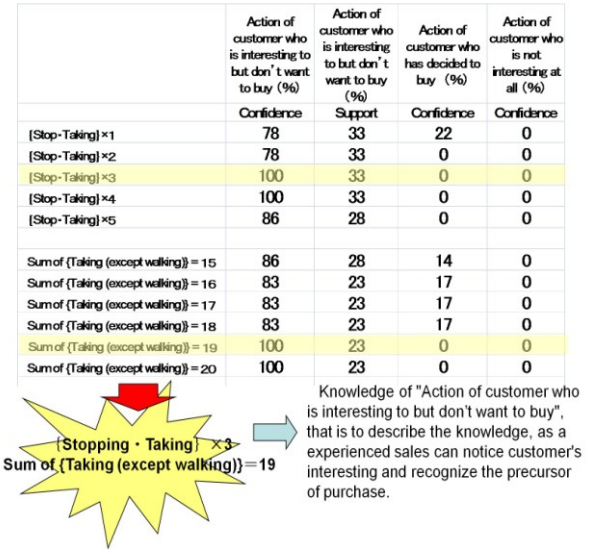


Fig.8. The knowledge is found by using association rule.

The discovered knowledge of predicting the shopping behavior can be applied in the robot system which based on Kukanchi and experimented on the environment as Fig. 1. The visible type robot provides the services to the customer for encouraging his purchase only when the customer is detected the characteristic behaviors which point to the action of interesting to the merchandise but don't want to buy.

V. THE SECOND EXPERIMENT: THE ROBOT SYSTEM OF THE INTELLIGENT ROOM

In the chapter, the authors assume a room as illustration in Fig.9, which is called intelligent room. It help the authors to explain the ideal of applying the knowledge by the mining technique which is mining in time-series motion logs, and show the possibility of using interactive data mining.

The intelligent room is assumed to acquire the appropriate knowledge from elders who are living alone by the normal living information which is collected from the intelligent room and use the interactive data mining to help predict the abnormal condition to assist the elders when they need.

The intelligent room set the high robust stereo cameras and the sensors that are installed in the doors of bathroom, toilet, entrance or any other place for detecting whether user is using it or not. The information which is collected from the room can be thought of the useful living information. Moreover, for directly interacting or supporting elders and

collecting information from elders' daily life, the authors also design to set an assistant robot in the room.

In the experiment, the visualization techniques are used as the first step of obtaining the knowledge that help to predicate abnormal health conditions of elderly people by using the detectable time-serious motion logs from the intelligent room.

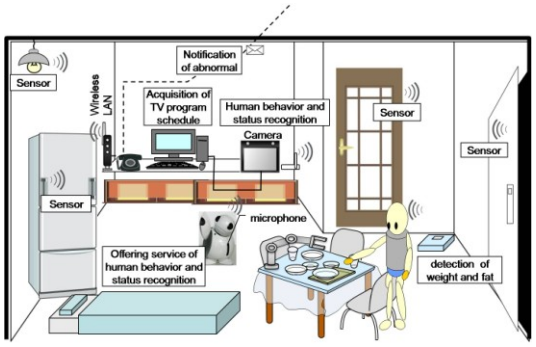


Fig.9. The intelligent room.

A. *Assistant robot*

For supporting the daily life of elderly people, the assistant robots which set in the intelligent room have to understand the situation and recognize the behaviors of users. Moreover, for directly supporting and interactive with the elders, the robot may include the function to communicate with people.

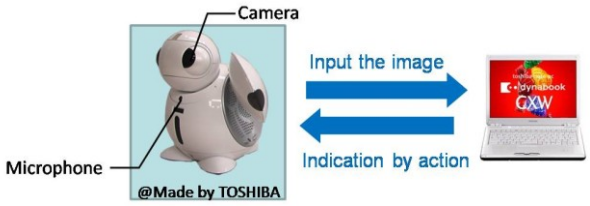


Fig.10. The robot which is named "ApriPoco". It is made by TOSHIBA and focus to communicate with human.

Fig.11. The image processing steps of extracting human motions.

For successful be the role of assistant robot, the authors choose "ApriPoco" as Fig.10, to be the assistant robot in the environment of the intelligent room. And the authors use OpenCV as the software to implement image processing for recognizing human motions. First, the camera captures the images by ApriPoco, and the system uses background subtraction technique, which is shown in Fig.11, to separate the target object and the background. Next, for recognizing human motions, the system extracts the three largest color areas of human skin and uses the extracted data to calculate the coordinates of the weight point of human skin area from the images. And then, the human motions can be detected by using the trajectory of coordinates of the weight point. By detecting the human motions, the robot can collect the direct information for processing data mining, and make communication with human for offering service.

As the authors described above, the robots have the functions to detect human motions and make a simple communication with humans. And for assistant elderly people's life, the robot may design to use those functions to achieve some tasks. For example, for maintaining elders' physical performance, the robots may have the function to improve elders' lifestyle by noticing their rhythm of life and supporting them does exercise appropriately. As reference [4], the motion of human can be used to calculate the amount of calorie. So, the technique can be use within ApriPoco and METs (Metabolic equivalents) which represents the intensity of physical activity, and with the weight of elder and the amount of time of doing exercise to calculate the amount of calorie. METs is the quotient of the amount of oxygen intake in a state divide by the amount of oxygen intake in the state of rest. The oxygen intake represents how many oxygen does a man need a unit of time. Because the amount maybe different from person to person, it may be referenced by the value which publics in ministry of health labour and welfare Japan as the standard value [5]. When elders doing exercise with ApriPoco, it captures the images and processes the images by the system. Then, the human motion is detected and the amount of calorie is also calculated. The amount of calorie is an important data for detecting the health status of elder people, too.

Furthermore, the robots for elderly people should not only for watching, but also can understand and interact with elders in psychology. So, the robots may have the functions to offer the regional information which fits to elder's hobby or interesting, and have the capability that is able to make a simple communication with elders.

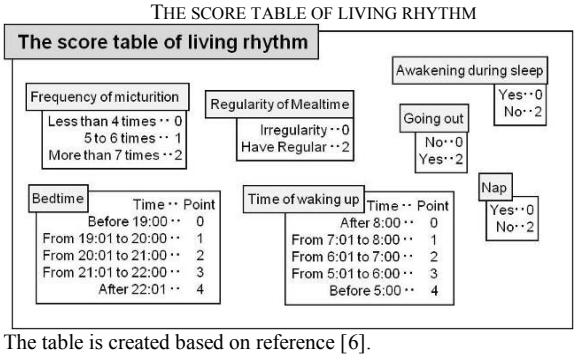
B. *The visualization technique for improving mining result*

The lifestyle shows the health status of a person. So, the authors design to install sensors and cameras into the intelligent room for taking down the elder's living information that can be thought of a useful data to know the elder's lifestyle and detect the health status of the elder.

Because the exactly data is not yet gotten, the authors based on the references [5] [6] to get the health status of three elders who are living alone to be an example for explaining. The points of living rhythm that represent the status of the elder's health are calculated based on the score table which is shown in Table IV. The table is created based on reference [6]. And the authors use METs [5] as the unit to calculate the amount of activities. As Fig.12, the authors can collect the data of living patterns of elders by their rhythm of

life and use the living patterns to know the health status of elders. Furthermore, the authors can implement the mining technique that is described in chapter 4 to distinguish the different situations of elders and find whom is prone to an abnormal status of health.

TABLE IV
THE SCORE TABLE OF LIVING RHYTHM



The table is created based on reference [6].

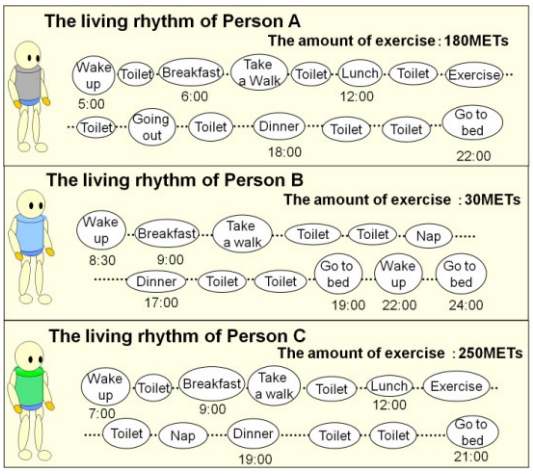


Fig.12. The lifestyle of three elders.

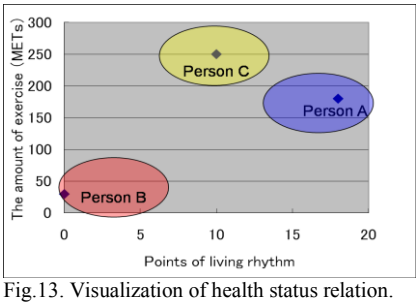


Fig.13. Visualization of health status relation.

The data of elders' living patterns that the authors described above can be illustrated as Fig.13. The X-axis in the graph shows the points of living rhythm and Y-axis shows the amount of exercise in a day by METs. The graph shows the visualized data of health status of the three elders from references. In Fig.13, person A shows the range of good health status that his health status is also be included, person B shows the range of not health status which is like him, and person C shows the range of who does exercise too much. The graph is not only easy to be understood by human, but also easy to find the relationship of every variable. In other words, the graph shows the usability of visualization techniques. The process of visualization techniques can be thought of the first step of knowledge discovery. The visualization techniques help people make judgments and decisions much appropriately. Adding the opinions of human is the target of performing interactive data mining which is useful for improving the performance of data mining and the accuracy of mining results.

By the intelligent room what the authors are assumed, the possibility of getting appropriate knowledge by using interactive data mining based on the information, which collected by sensors, cameras and the assistant robot in the intelligent room, is present. The knowledge for helping elders can be discovered by using the mining technique which the authors described in chapter 4, and the results of mining can be improved by using interactive data mining which includes the visualization techniques for supporting human make judgments.

VI. CONCLUSION

In this paper, the authors show two experiments which are constructed based on "Kukanchi". The first one shows the application of service robot system in a clothes store. And the other shows the system of elder supporting system. These experiments prove the possibility and utility of using the concept of Kukanchi to construct a system, recognize human action by mining in time-series motion logs, and the mining results can be improved by using interactive data mining which includes the visualization techniques.

ACKNOWLEDGMENT

This research is being conducted as a part of the "Promotion of Utilization of Skills to Support the Elderly" consigned by Bureau of Social Welfare and Public Health, Tokyo Metropolitan Government.

REFERENCES

[1] K. Shimokura, H. Tezuka, N. Katafuchi, T. Machino, Y. Nanjo, S. Iwaki, "Toward the Realization of Networked robot --Study of information distribution platform architecture for robot --",Technical report of IEICE,The Institute of Electronics Information and Communication Engineers.

[2] E. Sato, T. Yamaguchi, F. Harashima, "Natural interface using pointing behavior for human-robot gestural interaction", IEEE transactions on industrial electronics, 54(2), 2007, 1105-1112.

[3] T. Okada, H. Motoda, "Association rules and its surrounding." The Operations Research Society of Japan.

[4] T. Miki, E. Sato-Shimokawara, T. Yamaguchi, "Estimation the amount of exercise based on image processing by using communication robots." thesis of Tokyo Metropolitan University (2009 02)

[5] Ministry of Health Labour and Welfare Japan, "Exercise guidelines for health promotion 2006 -Prevention lifestyle related disease.(Exercise Guide 2006)"

[6] T. Ishikawa, T. Yuasa, Y. Motohashi, "Biosocial rhythm and living conditions of elderly people living alone in Akita city, in comparison with elderly people in shared living", Bulletin of School of Health Sciences Akita University,14(2), pp47-53(2006 10/31)