2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Generating Human Motion Transition Map in Indoor Environment and Analyzing Human Behavior by Geographical Clustering*

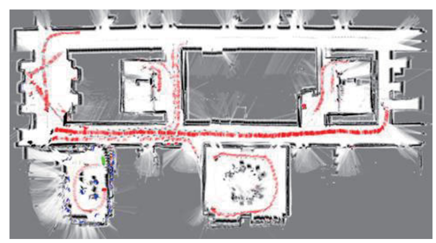
Yuji Ogawa, Zhidong Wang, *IEEE Member*, Tetsuya Wada, Yasuhisa Hirata, *IEEE Member*, and Kazuhiro Kosuge, *IEEE Fellow*

***Abstract*— In recent years, robots working in human living space with human-robot interactions are actively studied. To these robots, it is important to perform environmental cognition not only building environment map for autonomous motion of the robots but also estimating presences of human around the robots. In this study, by utilizing human state estimation function and SLAM based mapping technology, a concept and architecture of *Human Motion Map* by representing human behavior in the human living space as a hybrid map system are proposed. Beyond the conventional map which represents the existence of wall and objects, *Human Motion Map* represents not only the existence of humans in a particular location but also motion distributions. With recent improvements of the cloud computing technology, *Human Motion Map* can be accumulated as a kind of big data while measurements of robots are performed continuingly while it is moving around. In this paper, we propose a motion feature classification algorithm for clustering human motions geographically. Some experiment result of basic motion feature extraction, geographical clustering, and human motion behavior analyzing are provided for illustrating the validity of proposed algorithm**

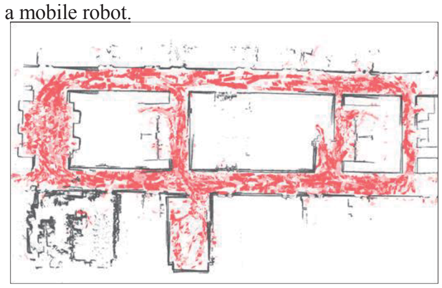
I. Introduction

In recent, rapid improvements on sensing and computational technologies have made very attractive researches on planning and mobile robot controls in human environments. Performing not only environment reconstruction such as map building but also human motion and motion pattern recognitions have to be taken in account when designing a robot. On environment mapping, Simultaneous Localization and Mapping (SLAM) technologies have been published in the last decade to solve the navigation and unknown environment motion problems[1]. On acquisition of human motion, Thompson, et al. developed a probabilistic strategy to build a human trajectory map on a known indoor environment based on both observations by multi-LRF sensors fixed in the environment [2]. Kanada et al. defined human behavior based on a map generated from observed human walking trajectories[3]. Katabira et al. proposed a method for analyzing crowds flow characteristics from LRF observation [4]. Mozos and Kurazume proposed a method for human detection and motion pattern analyzing with multiple 2D LRF sensors[5]. Song and Zha's group developed an on-line learning method to fusion vision and laser sensing data for tracking multiple walking targets [6]. Challaghan et al. proposed an algorithm to generate navigation maps based on human motion patterns [7].

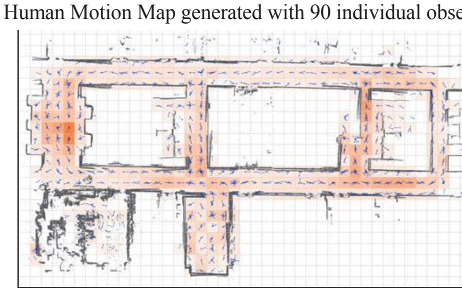
This research focuses on the construction of both the environment map and map of human motions simultaneously for mobile robots in indoor environments. By extending the mapping function of SLAM, we proposed the concept and architecture of Human Motion Map, representing extracted human behavior in the human environment [8]. In [9] and [10], a structure of high-dimensional human motion map with multiple human states is proposed and an algorithm implementing both human state estimation and human motion map update are introduced. Fig.1 shows results of human motion map generated from a large amount of individual observations, and consisting of various sub-maps for representing basic human walking behavior.



(a) A basic Human Motion Map represents various human motions observed by a mobile robot.



(b) A Human Motion Map generated with 90 individual observations



(c)A Human Motion Map representing statistical characteristics of human walking behivor

Fig 1. Different from most environment maps, Human Motion Map is designed as architecture of high-dimensional probabilistic map structure consisting a set of sub-maps.

* This work was supported in part by Grant-in-Aid for Scientific Research (B)(20360120), Japan.

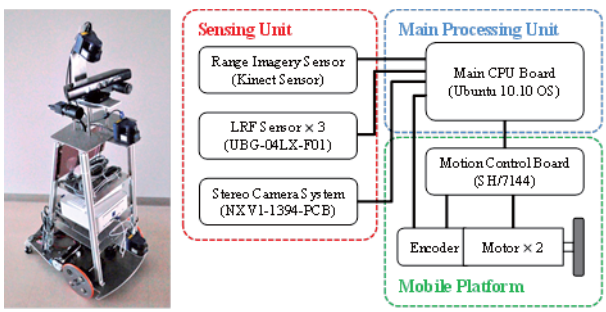
Y.Ogawa, Z.D.Wang, and T.Wada are with the Biomemtics Lab, Department of Advance Robotics, Chiba Institute of Technology, Japan.

Y. Hirata and K. Kosuge are with the System Robotics Lab., Tohoku University, Japan.

In this paper, it is proposed the concept of Human Motion Transition Map as a sub-map of Human Motion Map for representing and analyzing human behavior. A HMM based human action status estimation algorithm with continuous observation data are developed to generate the Human Motion Transition Map. Also, the geographical clustering method for analyzing Human Motion Transition Map is proposed, and some experiment results are provided to illustrate the validity of proposed method.

II. MANAGER ROBOT SYSTEM

In this research, a manager robot is developed for building environment map and performing object transportation tasks cooperatively with three worker robots and human operator.



(a) Manager Robot (b) Manager Robot System

Fig..2 The Manager Robot and its System Configuration

The robot consists of a two-wheel driven mobile platform with odometry function, sensing unit for information acquisition of human and environment, and main processing unit for map generations and its cooperative task planning (Fig.2). Three LRF sensors are used for both environment mapping and human estimation. Two LRF sensors are installed with horizontal scanning planes at the 0.8m and 0.4m height positions respectively for detecting upper bodies and legs of humans. By fusing the information from two sensors (Fig.3), higher reliability on human detection and map building are achieved [9]. Using the OpenNI Library[11], a Kinect sensor is equipped for tracking 3D motion of humans in front of the robot and estimating human posture and states (Fig.3). Another LRF is installed in the top of the robot and its scanning plane is set as -30 degree for obtaining scanning data of surfaces of tables and chairs in the environment. The ROS Gmapping package[12][13] developed by Brian Gerkey, et.al., is used for environment map building and mobile robot's localization.

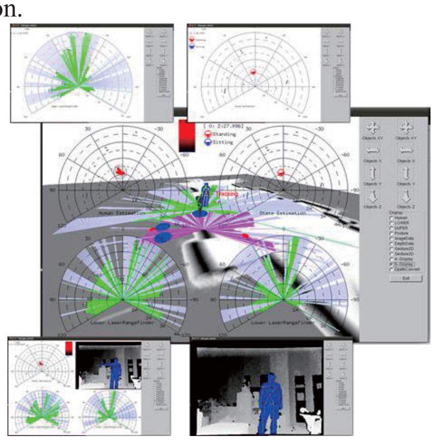


Fig.3 GUI of Human Existence and State Estimation System

III. HUMAN MOTION STATUS ESTIMATION

In the same manner as localization and mapping, a LRF sensor shows good performance to track moving objects including human beings in a relatively wide area. On the other hand, a Kinect sensor is powerful on observing human gesture and other human body motions. In this research, a human motion status estimator which consists of two estimation components, lower body motion estimator and upper body action estimator, is developed. In processing sequential observation data, the lower body motion estimator is designed to extract motion transition between various lower body motions of people based on LRF scanning data. The upper body action estimator is designed for estimating an action itself, because in general, our upper body actions are either gestures for communication or task instruction, or actions for object handling. Both estimators are implemented by using Hidden Markov Model (HMM) for modeling and estimating sequential motion of human. By applying Multidimensional Scaling (MDS), two multi-dimensional symbol spaces are generated from results of HMM training, and Kullback-Leibler divergence is used for calculating similarities of observed actions of human with predefined human gestures and motion transition patterns of upper and lower body respectively.

A. Estimation of Upper Body Action

Estimations of upper body action of a person are based on processing human action sequence by using the HMM developed. Since a gesture of a person can start at any moment which the robot is not able to know in advance, one solution is processing all action sequence data started from all moments, and using the result with high similarity scores. This will be time consuming for a real time estimation procedure which is running on the robot's on-board CPU. We notice that gestures of a human for communication or task instruction are usually starting from an action to arise his/her hand from lowest position, and use this as the starting trigger for extracting motion data sequence.

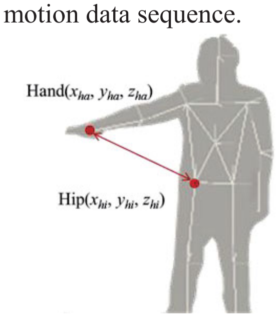


Fig.4 Skeleton Model Hand Action for Calculating the Starting Trigger

The starting trigger is set as when d_h, Euclidean distance between hand (x_{ha}, y_{ha}, z_{ha}) and nearby tip point (x_{hi}, y_{hi},z_{hi}) calculated from skeleton tracking data of Kinect sensor (Fig.4), is larger than 280[mm]. Fig.5 shows GUI of upper body gesture estimator developed which consists of camera image, skeleton image and visualization result of symbol space of action estimation at the upper part of the window, and X,Y and Z components of velocities of human's hand at the lower part of the window. The hand velocity data are plotted with light blue background when the system senses the starting trigger, arising up a hand, and upper body motion data observed by Kinect sensor are extracted as the data sequence for HMM motion estimator. The data sequence is terminated when the hand of the observed person is put down or 30 frame data are obtained.

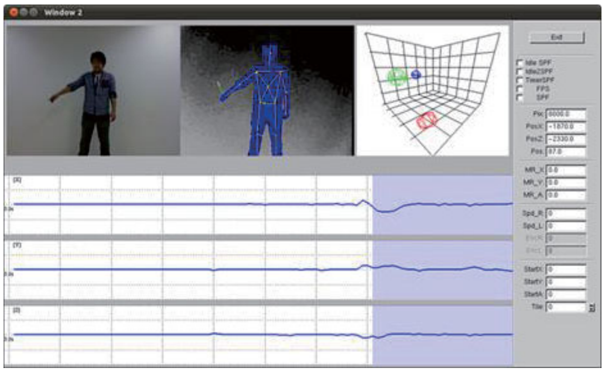


Fig.5 GUI Window Developed for Upper Body Gesture Observation and Estimation

In the observed data vector O[t] = [o_1; o_2; o_3; o_4; o_5; o_6], o_1 denotes facing angle of tracked skeleton, o_2 and o_3 denote angle of shoulder and elbow joint respectively, o_4, o_5 and o_6 are X, Y and Z components of hand speed respectively. Information of the facing angle is used for estimating the degree of human intention on communicating to the robot. It is easy to know that shoulder and elbow joint angles are tightly related with representations of human gestures for communication. Also both angles are able to be obtained from skeleton tracking data. Finally, the data set of 3D vector velocity of human's hand is one of the most important parameters for identifying human gestures in communication or actions on performing some tasks.

For estimating human intention on commanding and controlling robots, five gesture actions: *Pointing Position(λ_1)*, *Trajectory Pointing(λ_2)*, *Beckoning(λ_3)*, *Retire(λ_4)*, and *Stop(λ_5)* are defined and implemented in the upper body action estimator. Table I shows results of HMM training for these five basic actions and other five actions as a control group. As the result, to top five basic actions, symbol values log(O/λ) of the main diagonal are in order of 23, and other values are larger than 100. To other five actions different from trained gestures, symbol values are over 100 also.

Table 1 Result of log(O|λ) for HMM based Upper Body Gesture Estimation

	λ_1	λ_2	λ_3	λ_4	λ_5
O_1: Pointing Position	23.3	118.0	321.4	211.3	167.8
O_2: Trajectory Pointing	118.0	23.2	233.1	301.1	229.5
O_3: Beckoning	247.0	153.1	23.1	251.6	310.1
O_4: Retire	258.0	375.0	272.8	23.2	393.1
O_5: Stop	226.3	183.8	80.2	221.0	23.2
O_6: Waving Hand (S)	193.6	133.5	93.7	215.0	232.4
O_7: Waving Hand (L)	294.6	228.6	240.0	224.1	210.3
O_8: Deskwork	111.7	100.7	175.3	223.4	126.1
O_9: Picking Object	300.7	369.1	437.3	343.4	352.2
O_{10}: Putting Object	198.8	143.8	151.8	190.4	209.5

Fig.6 shows the five upper body gestures in 3 dimension symbol space: *Pointing Position (large blue ball), Beckoning(large green ball), and Stop(large green ball)*, *Trajectory Pointing(middle size blue ball)*, *Retire(middle size yellow ball)*, as well as other untrained action represented as five small balls. In the symbol space, the first dimension corresponds to the element of low speed of hand motion, the second dimension corresponds to the element of high speed of hand motion, and the third dimension corresponds to the element of the angle of human's hand.

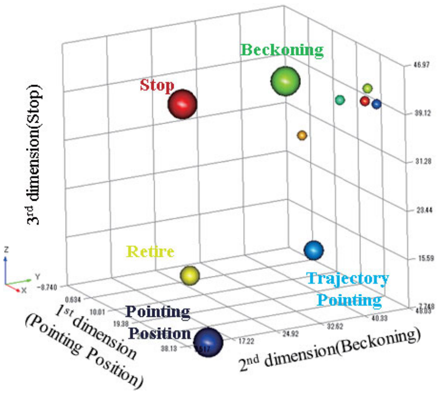


Fig.6 Symbol Space of Upper Body Gesture

B. Estimation of Lower Body Motions based on Motion Transition

Observations of lower body motions are performed by using two LRF sensors, the sampling rate is 50ms. The motion data is generated from three scans of target subjects with a tracking algorithm implemented, and motion data is obtained in every 150ms and is store in the human motion map at every 500ms. Fig.7 is a visualization result of lower body observation.

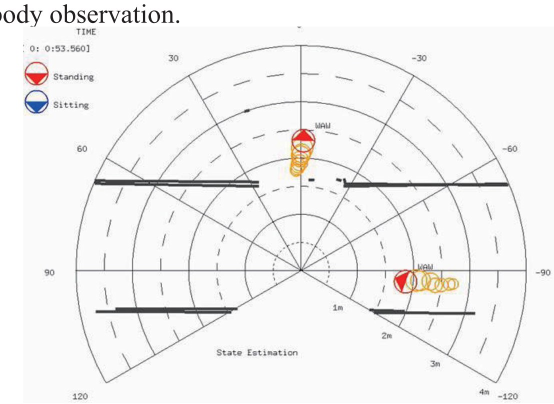


Fig. 7 Example of Observation Result of Lower Body Motion

The observation data vector of lower body O[t] = [o_1; o_2] consists of two symbol elements: moving distance of the lower body observed (o_1), and relative distance between body and leg which is used for checking the sitting posture (o_2). In the estimator of lower body motion, body motion patterns are defined as three basic motions (*Walking, Standing* and *Sitting*) and 6 motion transition patterns between two of three basic motion elements. Table 2 shows the HMM training result of motion transition of lower body. As a result, symbol values of the main diagonal are in order of 7, and other values are larger. Within them, 6 values are in value around 8 and 9 which indicates similar motion pattern in the symbol space, and many other are over value of 100.

Table 2 Result of log(O|λ) for HMM based Lower Body Motion Estimation

	Motion Pattern	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9
$\boldsymbol{O}_1$	Walking→Walking	7.2	32.5	239.4	27.0	9.3	100.2	229.6	8.1	9.5
$\boldsymbol{O}_2$	Standing→Standing	28.4	7.2	159.8	11.7	19.4	47.8	173.2	18.4	20.6
$\boldsymbol{O}_3$	Sitting→Sitting	223.6	161.3	7.1	214.2	201.3	16.0	11.7	219.6	11.7
$\boldsymbol{O}_4$	Standing→Walking	24.2	16.4	230.9	7.7	18.1	88.5	224.6	12.9	18.8
$\boldsymbol{O}_5$	Walking→Standing	24.6	18.7	207.1	18.5	7.7	73.6	198.8	13.3	18.2
$\boldsymbol{O}_6$	Standing→Sitting	306.2	249.6	21.8	316.0	282.1	7.6	11. 1	307.6	16.3
$\boldsymbol{O}_7$	Sitting→Standing	278.0	230.2	37.6	289.1	256.3	78.3	7.6	285.6	45.5
$\boldsymbol{O}_8$	Walking→Sitting	304.2	250.3	218.0	318.8	281.5	81.9	223.4	7.6	230.2
$\boldsymbol{O}_9$	Sitting→Walking	8.4	22.1	221.8	18.5	8.1	86.3	198.8	9.0	7.5

Fig.8 shows the lower body motion pattern in 3 dimension symbol space. Three main basic motions are shown as large balls: red ball (*Sitting*), green ball (*Standing*) and blue ball (*Walking*) respectively, and other 6 small balls represent other motion transition patterns in the symbol space. In this three dimension symbol space, three basic motion symbols are located in linearly relation, and 6 small balls are around these three main motion too. These provide use a relatively good performance on motion estimations of lower body motions as well as upper body gestures.

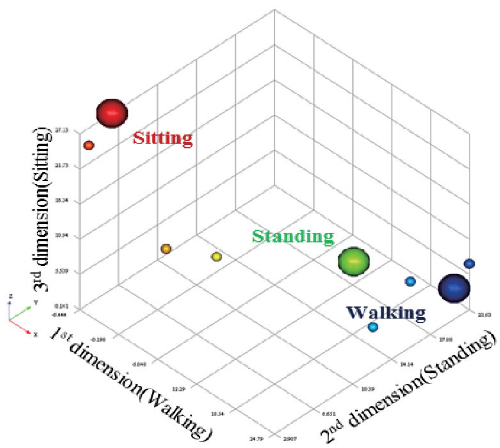


Fig. 8 Symbol Space Construction base on Lower Body

IV. Generating Human Motion Transition Map

We incorporate the human gesture estimation and motion transition estimation into the human motion map structure that we proposed [8-10]. To general environment maps generated SLAM and other mapping method, a map can be with a relatively simple structure which only need to represent the probability of the walls or objects in each location. Different from environment maps, the proposed human motion map is a high-dimensional hybrid map system consisting of multiple sub-maps for recording and representing different human states observed respectively. We generate a set of sub-maps with data of motion transition for recording and representing where each transition pattern occurs, and name it *Human Motion Transition Map*.

Similar with *Human Motion Map*, *Human Motion Transition Map* is accumulated from many observation experiments, and all motions and motion transition data are recorded together with environment map. Therefore, not limited to obtaining a motion pattern of a particular subject, *Human Motion Map* and *Motion Transition Map* are databases which are able to be used for obtaining general behaviors of human motion, especially behaviors tightly related with particular objects or locations. Because *Human Motion Map* and *Motion Transition Map* are as large as indoor environment area that the robot explores and maps via SLAM algorithm, we use computational clustering and labeling method for obtaining human behaviors from *Human Motion Map*s.

In this session, a transition map including *Walking*, *Standing* and transition pattern between these two motions is provided, and the analysis based on geographical clustering of mapping data are demonstrated to show the potential of *Human Motion Map* on obtaining human behavior.

A. Generating Human Motion Transition Map in Font of a Room

As an example, a set of human motion data in front of a room is obtained in an experiment setting shown in Fig.9-(a). Several subjects are walking into a door of the room or passing through the corridor, and the scanning data are shown in Fig.9-(b).

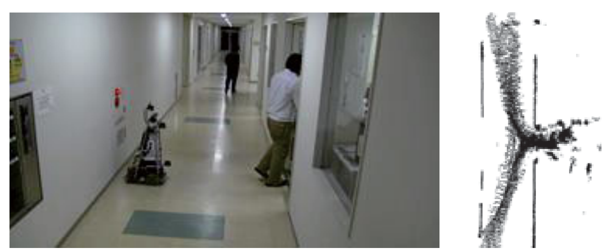
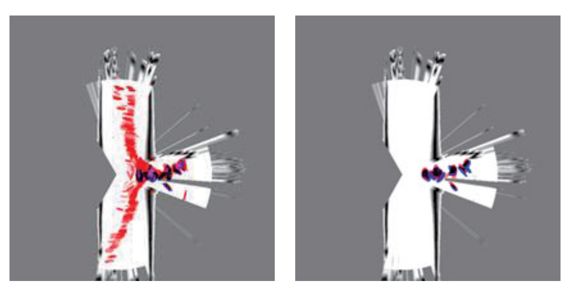


Fig. 9 (a) Experiment Environment(left) (b) Scans Data(right)



(a) Human Motion Map (b) Human Motion Transition Map

Fig. 10 Human Motion Map and Motion Transition Map Between Walking and Standing States

Fig.10-(a) is a sub map of Human Motion Map which representing the two basic motions: red denotes walking and blue denotes standing. Fig.10-(b) is a Human Motion

Transition Map which denotes motion transition from *Walking* to *Standing* in blue, and motion transition from *Standing* to *Walking* in red respectively.

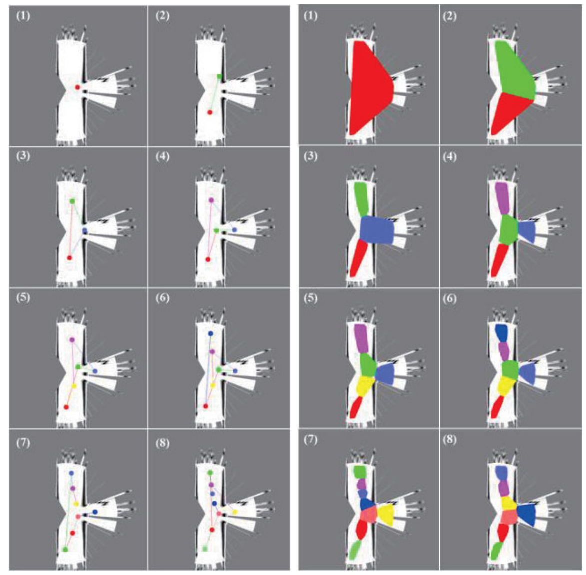
From Human Motion Map data, a geographical trajectory pattern of straight walking in a corridor and a trajectory pattern connecting from two directions in front of the door are easily been confirmed. Also from Human Motion Transition Map, we can easily confirm that motion transition between Walking and Standing happen in front of and in the door way. This indicates the motion patterns of body motion when subjects are opening a door and changing in-room shoes. This kind of motion behavior is related strongly with objects in particular location such as, doors, tables and chairs.

B. Geographical Clustering and Human Motion Behavior Analyzing

For analyzing human motion behaviors and geographical relationship, we developed a *K*-means method based geographical clustering package and implemented into the Human Motion Mapping system. The geographical clustering algorithm is applied to *Human Motion Map* and *Motion Transition Map* data generated in the previous sub-session. The results of clustering in 1-8 clusters are shown as different color regions and center positions of clusters in Fig.11 and Fig.12. In Fig.11-(a), lines connected between centers of neighbor clusters shows walking direction. However, at the place near the door, lines with big direction changing angle show feasible walking direction, lines with relatively small direction changing angle are not feasible, and cannot be used for extracting walking behaviors. Results of Fig.11-(b) indicate that regions of clusters with small clusters numbers (3 in this map) may overlap some areas where people cannot walk through, and regions of clusters with a certain large number (4 and more clusters) will represent walking regions correctly. Also it can be confirmed that sharp and size of the cluster's regions in corridor change much when the cluster numbers are change from 4 to 8, but clusters near the door are not showing so much changes. Fig.12 shows not only that motion transitions between *Walking* and *Standing* motions mostly are observed near the door, but also two transitions: *Walking to Standing*, and *Standing to Walking*, are both with two regions with center positions are located alternatively on the walking direction. This result of geographical clustering indicates that human behavior near the door usually switch their walking motion for opening a door then changing their shoes to in-room shoes, a "Japanese life style". Therefore, as showing by this example, many human behaviors at particular locations or around some particular objects can be extracted and labeled by applying a computational geographical clustering algorithm to Human Motion Map and Human Motion Transition Map when we collected enough data.

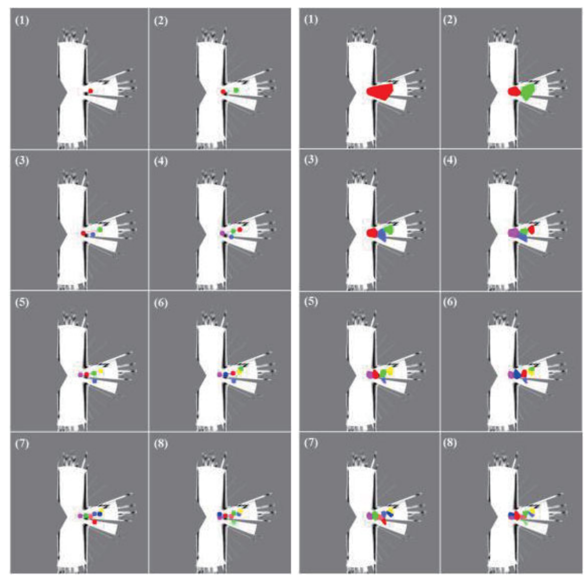
C. Experiments in Various Indoor Environments

Several experiments were conducted in the 18th floor of the Building 2 at the CIT (Fig.13), and we especially paid attention to obtain motion data in 7 areas where environments have different characteristics and persons may show different motion behaviors. Multiple subjects were showing daily gestures in these areas.



(a) Center Positions of Clusters *(b) Regions of Clusters*

Fig. 11 Results of Geographical Clustering of Human Motion Map



(a) Center Positions of Clusters *(b) Regions of Clusters*

Fig. 12 Result of Geographical Clustering of Motion Transition Map

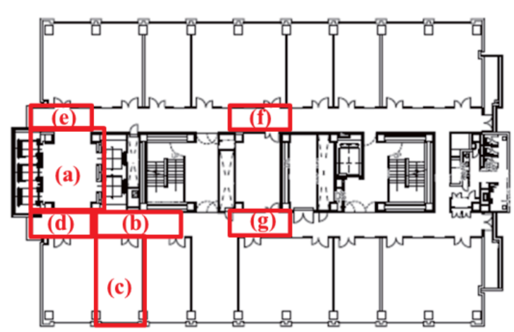
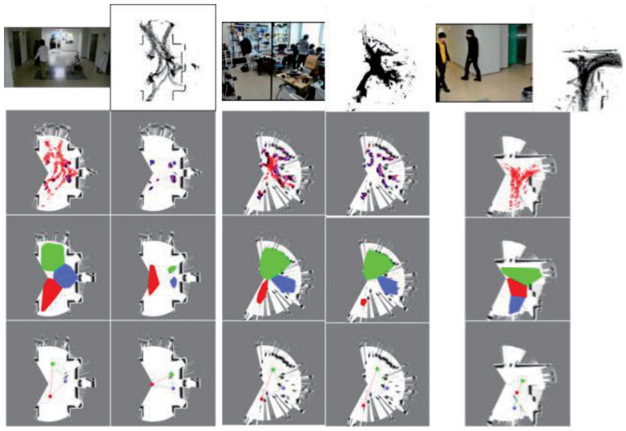


Fig. 13 Experiment Environment at 18th Floor of CIT Building 2



(a) (c) (d)

Fig. 11 Results of Geographical Clustering of Human Motion Map and Human Motion Transition Map in Various Environments (a) Elevator Hall (c) Lab's Room (d)a Corner from Corridor to Elevator Hall

In the area (a), Elevator Hall, human's motions have the same purpose, to wait and take an elevator, and their motion trajectories usually are connected to several particular locations of elevator doors. As the results shown in Fig.11-(a), clusters from Human Motion Map data indicates waiting actions in front of different elevators clearly, and clusters from Human Motion Transition Map data also show another basic human motion behavior in elevator halls: motion transitions happened for pushing elevator buttons beside the elevators. Results for area (b) were addressed previously. Area (c) is the Lab's room and there are much more objects around and human's motions are not so simple as other areas, but clustering results still show some basic behaviors, such as avoiding obstacles and showing motion transitions in front of desks, chairs and the door (Fig.11-(c)). Fig.11-(d) shows results of walking behavior at the Corner from the corridor to the elevator hall (area (d), as well as area (e)). With the exception of some rare Standing cases, people work at certain speed in the corridor. After passing over the corner of the corridor, walking trajectories are spread to two directions for approaching two sides of the elevator hall where elevator buttons are set. This kind of walking trajectory pattern is useful for motion planning and control of avoiding invisible person when a robot is moving near the corner and is with limited visibility. In area (f) and (g), human motion behavior is simple, straight walking with some obstacle or human avoidance trajectory. From these results, several basic human behaviors with geophysical dependency were obtained by computational cluster of the *Human Motion Transition Map* proposed.

V. CONCLUSION

In this paper, we proposed a concept of Human Motion Transition Map as a sub-map of Human Motion Map for representing and analyzing human behavior. A HMM based human action status estimation algorithm with continuous observation data are developed by establishing symbol spaces from known basic motions and calculating similarities. By incorporating the motion estimation algorithm to a developed SLAM based mapping system, Human Motion Map and Human Motion Transition Map are generated. The geographical clustering method for analyzing Human Motion Transition Map is proposed to analyze and investigate the human motion behaviors with geographical dependency. Some experiment results are provided to illustrate the validity of the proposed method, and showing basic motion behaviors extracted in the indoor environment. Investigation Human Motion Map system with large amount data, and developing frameworks and algorithms to apply extracted human motion behavior information to motion control of mobile robots are the future works of this research.

REFERENCES

[1] S.Thrun, W.Burgard and D.Fox, "Probabilistic Robotics", The MIT Press, 2005.

[2] S.Thompson, T.Horiuchi and S.Kagami, "A Probabilistic Model of Human Motion and Navigation Intent for Mobile Robot Path Planning", Int. Conf. on Autonomous Robots and Agents(ICARA09), pp.663-668, 2009.

[3] T.Kanda, Dylan F.Glas, M.Shiomi, H.Ishiguro and N.Hagita, "Who will be the customer?: A social robot that anticipates people's behavior from their trajectories", Tenth Int. Conf. on Ubiquitous Computing (UbiComp 08) , pp.380-389, 2008.

[4] K.Katabira, T.Suzuki, H. Zhao, Y.Nakagawa, R.Shibasaki, "An Analysis of the Crowds Flow Characteristics by using Laser Range Scanners", Proc. of Asian Conference on Remote Sensing (ACRS07), pp.953-956, 2007.

[5] O.M.Mozos, R.Kurazume and T. Hasegawa, "Multi-part people detection using 2D range data", International Journal of Social Robotics, Vol. 2, No. 1, pp. 31-40, 2010.

[6] X. Song, H.Zhao, J.Cui, X. Shao, R.Shibasaki, and H.Zha, "Fusion of Laser and Vision for Multiple Targets Tracking via On-line Learning", Proc.2011 IEEE Int. Conf. on Robotics and Automation(ICRA11), pp. 4333-4340, 2011

[7] Simon T. O'Callaghan, Surya P. N. Singh, Alen Alempijevic and Fabio T. Ramos, "Learning navigational maps by observing human motion patterns", Proc.2011 IEEE Int. Conf. on Robotics and Automation(ICRA11), pp. 4333-4340, 2011

[8] T.Wada, Z.Wang, T.Matsuo, Y.Ogawa, Y.Hayashibara, Y.Hirata and K.Kosuge, "Building Human Motion Map for Mobile Robot in the Indoor Dynamic Environment", IEEE Int. Conf. on Robotics and Biomimetics(ROBIO10), pp.543-548, 2010.

[9] Y.Ogawa, Z.Wang, T.Wada, K.Tomiyama, Y.Hirata and K.Kosuge, "Building Human Motion Map for Mobile Robot in the Indoor Dynamic Environment", IEEE Int. Conf. on Robotics and Biomimetics(ROBIO11), pp.543-548, 2011.

[10] T.Wada, Z.Wang, Y.Ogawa, Y.Hayashibara, Y.Hirata and K.Kosuge, "Incremental Human Motion Map System and Human Walking Behavior Representation in Indoor Environment", IEEE Int. Conf. on Robotics and Biomimetics(ROBIO12), pp.747-752, 2012.

[11] The OpenNI Organization, "Introducing OpenNI", Open Natural Interaction Library, http://www.openni.org, (accessed 2010-12-8).

[12] Brian Gerkey, Cyrill Stachniss, and Wolfram Burgard, "https://www.openslam.org/gmapping.html, (accessed 2014-2-10).

[13] Brian Gerkey, Cyrill Stachniss, and Wolfram Burgard," Improved Techniques for Grid Mapping with Rao-Blackwellized Particle Filters", IEEE Transaction on Robotics, Volume 23, pages 34-46, 2007