35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Human Body Contour Data Based Activity Recognition

Nergui Myagmarbayar, Yoshida Yuki, Nevrez Imamoglu, Jose Gonzalez, Mihoko Otake and Wenwei Yu, *Member*, *IEEE*

Abstract— **This research work is aimed to develop autonomous bio-monitoring mobile robots, which are capable of tracking and measuring patients' motions, recognizing the patients' behavior based on observation data, and providing calling for medical personnel in emergency situations in home environment. The robots to be developed will bring about cost-effective, safe and easier at-home rehabilitation to most motor-function impaired patients (MIPs). In our previous research, a full framework was established towards this research goal. In this research, we aimed at improving the human activity recognition by using contour data of the tracked human subject extracted from the depth images as the signal source, instead of the lower limb joint angle data used in the previous research, which are more likely to be affected by the motion of the robot and human subjects. Several geometric parameters, such as, the ratio of height to weight of the tracked human subject, and distance (pixels) between centroid points of upper and lower parts of human body, were calculated from the contour data, and used as the features for the activity recognition. A Hidden Markov Model (HMM) is employed to classify different human activities from the features. Experimental results showed that the human activity recognition could be achieved with a high correct rate.**

I. Introduction

Recently, due to increasing elderly, the improvement of medical treatment and prevention of lifestyle diseases, home healthcare is more on demand. The aim of the home healthcare is to reduce hospital admissions and to make it possible for people to do rehabilitation exercise at home. Human motor function impairments are empirically observed from human motions and gait movements.

Human motion, such as gait, is usually measured by using motion capture systems. However, motion capture systems are costly and only effective in limited areas, thus not suitable for at-home monitoring. In this research, instead of the motion capture system that needs multiple cameras installed at fixed location, one camera used as a vision of the mobile robot was used to recognize the human activities. This requires the robot to track the human subject appropriately, and observe and analyze human activities.

Our research objective is to develop autonomous mobile home healthcare and rehabilitation robots for motor-function impaired patients (MIPs). The robot to be developed should be able to track patients' motion, accurately measure and recognize patients' activities in home environment.

In our previous study [1], [2], we used a mobile robot equipped with a Kinect sensor, which contains a RGB camera, a depth sensor with an infrared emitter, and multi-array microphones. From robot vision using the Kinect sensor, the mobile robot was controlled to track patients, and patient's activities were analyzed using the RGB-D images from the Kinect sensor [1], [2]. In [5], [6], depth cameras were used for indoor localization and navigation. However, the precise human measurement by a mobile robot, as well as activity recognition using measured data from a mobile robot has not been sufficiently studied. In our previous study [1], human lower limb joints (hips and knees) and trunk measured by the mobile robot were used to recognize human activities (standing, sitting, walking, impaired walking and falling down). However, the depth image data, and joint angle data extracted from the RGB-D image data, were unavoidably influenced by the motion of human subject, and the robot itself. Although the influence could be reduced by finding a better viewpoint for the robot, and improving the stability of robot motion, it is still difficult for the joint data measured by the mobile robot to be the ideal signal source for activity recognition. In order to get accurate joints angle, we used several color markers attached to lower limb joints, i.e., hips, knees and ankles joints. But color markers attached to joints could be inconvenient for subject in daily life.

That is the reason why, this study is aimed to enhance the human activity recognition by using different source data such as the contour data of the tracked human subject extracted from the depth images of the Kinect sensor located at the bio-monitoring mobile robot.

This study deals with the human activities (standing, sitting, bending, walking and lying down) recognition by using several geometric parameters calculated from the contour data of the tracked human subject, such as the ratio of height to weight of the tracked human subject, and distance (pixels) between centroid points of upper and lower parts of human body, while removing the color markers attached to lower limb joints. Hidden Markov Model (HMM) based human activity recognition taking the contour data of the tracked human subject extracted from depth image of Kinect sensor as input was proposed.

II. Related Work

There are many research works related to mobile robot tracking and following a subject [3-6]. They basically focused on only tracking and following the subject, and not for

*Research supported by JSPS Grants-in-Aid for Scientific Research #23・89, in part by Japanese Science Technology, Strategic International Research Cooperative Program (SICP), Japan-USA, FY2011-2013, Bio-monitoring robot project grant, and in part Grant-in-Aid for Scientific Research on priority area Founding a creative society via collaboration between humans and systems (#4101) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

All Authors are with Graduate School of Engineering, Chiba University, Chiba-shi, Inage-ku, Yayoi-cho 1-33, 263-8522 Japan. (corresponding author's e-mail: myagaa@ graduate.chiba-u.jp)

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

bio-monitoring patients. There is not any application of bio-monitoring mobile home healthcare system. There have been many research works related to human activities recognition and interpretation [7-16]. Most of them are using image processing techniques, such as binary image extraction, gait energy image, and so on. These studies have some limitations, such as, offline recognition, the camera is fixed and vision area is limited. Fixed and static video cameras used are recorded image sequences, and image processing procedures are followed by at first took background image capture, and then extracted human body from captured image sequences by subtracting the background image from image sequences. These studies are not suitable for dynamic environment and in real time. This study is aimed to bring human activity recognition system at any dynamic environment and in real time by dealing with the contour data of the tracked human subject extracted from the depth image of the Kinect sensor.

III. Framework of Human Activity Recognition

This study is aimed to overcome some weak points of our previous study [1], [2], such as influences of human motion and robot to lower limb joints angle, and using many color markers attached to joints. In this study, we extracted the contour data of the tracked human subject extracted from only the depth image to recognize human activities.

Fig. 1 shows a framework of human activity recognition from depth image of the Kinect sensor.

We can get two kinds of images from the Kinect sensor, such as RGB image and depth image shown in Fig. 1. In this study, we used only depth image for getting human motion information data. Image processing and activity recognition procedure of our framework are explained in the following sections.

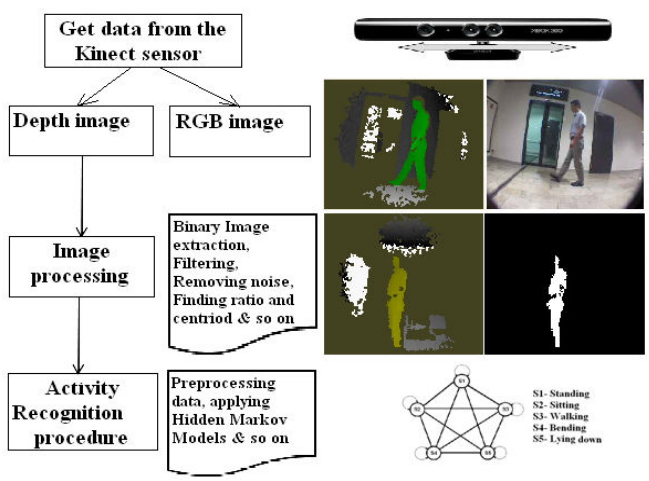

Fig. 1. A framework of human activity recognition from depth image of the Kinect sensor.

A. Image processing

In our image processing technique, we first extracted binary image, which includes the contour data of the tracked human subject extracted from the depth image captured from the Kinect sensor located at the bio-monitoring mobile robot. In binary image extraction, in order to extract the contour data of the tracked human from the depth image, we used likelihood data of human body shape, then used some image processing functions, such as filtering function for removing noises, image gap filling, extracting blobs from the binary image, and so on.

In order to get several geometric parameters from the contour data of the tracked human subject, we calculated the following data from the binary image represents the contour of the tracked human body.

1) *Ratio of height to weight of the contour data of the tracked human subject*
2) *Centroid points of upper part (one of third of human body shape) and lower part (one of third of human body shape)*
3) *Distance (pixels) between centroid points of upper part and lower part*

Fig.2 shows the method of finding some geometric parameters such as, the weight and the height of the contour data of the tracked human subject activities used in this study.

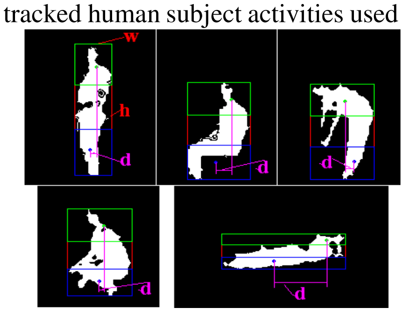

Fig. 2. Finding height (h) and weight (w) (red boxed area), and distance (pixels) between the centroid points of upper (green boxed area) and lower (blue boxed area) parts of binary image expresses the contour data of human body from side viewpoint

In Fig.2, *w* refers the weight of the contour of human body, *h* refers the height of human body shape, and *d* refers the distance (pixels) of these two centroid points. From this data, we found the ratio (*h/w*) of height to weight of the contour of human body in further usage of human activity recognition. The ratio of height to weight of the contour data of human body was enough to distinguish some human activities (standing, walking and lying down) used in this paper, but it was sometimes not enough to distinguish between sitting and bending type 2 activities. In order to differentiate these two human activities from each other, from the contour data of human body we extracted the centroid points of the upper part (green boxed area) and the lower part (blue boxed area) of the contour data of human body shown in Fig.2. Then, we calculated the distance (pixels) of these two centroid points, denoted by *d*.

B. Hidden Markov Models

Hidden Markov Model is a statistical model, in which the system processed is considered as a Markov process having unknown parameters. The application of Hidden Markov Model is to detect the sequence of states using observable data. We assumed that human activity could be expressed by five different states (S1-Standing, S2-Walking, S3-Sitting, S4-Bending (two kinds of bending activities shown in Fig. 4g, i) and S5-Lying down).

The HMM in this study was characterized by the following parameters.

- The number of states of the model, N=5
- The number of distinct observation symbols per state, M=7, quantization data of the ratio of height to weight of human body shape, and quantization data of difference distance (pixels) between the centroid points of upper and lower parts of human body shape
- The state transition probability distribution A={a_{ij}}

  $a_{ij} = P(q_t = S_j \mid q_{t-1} = S_i)$
- The observation symbol probability distribution in state j; $b_j(k) = P(V_k \ at \ t \mid q_t = S_j)$

where $V = \{v_1, v_2, ..., v_M\}$ are distinct observation symbols.

- The initial state distribution: $\pi_i = P(q_1 = S_i)$
- The model parameters notation: $\lambda = (A, B, \pi)$

C. Data preprocessing

Once, we calculated several geometric parameters from the contour of the tracked human subject, such as the ratio of height to weight of the contour of human body and the distance (pixels) between centroid points of upper and lower parts of the contour of human body from binary image, we preprocessed it for further analysis.

Before the process of HMM, several geometric parameters (the ratio of height to weight of the contour of human body and the distance (pixels) between centroid points of the upper and the lower parts of the contour of human body) of the counter of the tracked human subject were quantized by a set of threshold values to several discrete levels shown in Table I. The threshold value set was determined by trial-and-error.

TABLE I. QUANTIZATION OF ANGLE DATA

Name of Data	Threshold value	Quantization level
Ratio of height to weight of the contour of human body	>2.75	1
	2-2.75	2
	1.6-2	3
	0.8-1.6	4
	<0.8	5
Distance (pixels) between centroid points of upper and lower parts of human body	>30	6
	<30	7

We created two kinds of datasets from experiment data of each subject, such as, training dataset consisted of 3-5 seconds duration data from each activities, another dataset is test dataset, which consisted of remaining datasets.

IV. EXPERIMENTAL RESULTS

In the human activity recognition experiment, 3 subjects (1 female and 2 male) were recruited. The objective and procedure were explained, with informed consent signed, before the experiment. The subjects were asked to do the following activities: walking with their normal speed, standing up, sitting down, bending type 1, bending type 2, and lying down. The sequence of the activities made a trial, forming a data set. Each subject was required to perform activities individually shown in Table II. In addition, each subject was asked to perform whole dynamic sequential activities consist of all activities.

According to our framework of human activity recognition from depth image of the Kinect sensor, human body shape information data, such as the ratio of height to weight of human body shape, and difference distance between centroid points of upper and lower parts of human body shape, were calculated and preprocessed for human activity recognition by using Hidden Markov Models.

TABLE II. EXPERIMENT ACTIVITIES AND ITS DURATION

Activities	Walking	Standing	Sitting	Bending	Lying down
Experiment activity duration of each subject, and number of trials	10 sec 5 trials	5 sec 9 trials	5 sec 9 trials	5 sec 18 trials	3 sec 9 trials

HMM were built from one sequence of human body shape information data consisted of all states data. Different sequences of human body shape information data were performed to recognize the gait gesture using HMM and to find the most likely decoded paths (Viterbi paths).

For each subject, one data set was selected as training data set for estimating the model parameters, i.e., the state transition probability distribution, the observation symbol probability distribution and the initial state distribution. Then, we tested the model, with the other data sets, and calculated a correct recognition rate.

Fig. 3 shows one data set of individual activities and recognition results. Note, the sequence of the movement that the subjects were required to perform is: Walking (S2) – standing (S1) –sitting (S3) –standing (S1) –walking (S2) – bending type 1 (S4) –standing (S1) –bending type 2 (S4) – standing (S1)- lying down (S5), expressed as 2-1-3-1-2-4-1-4-1-5.

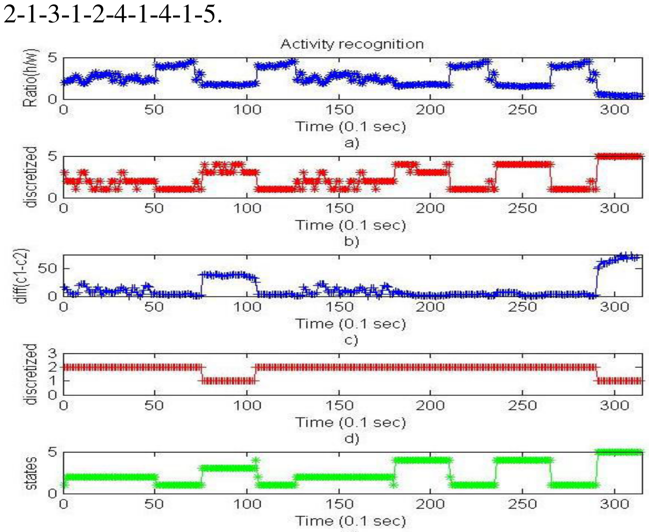

Fig. 3. Human activity recognition results. y axis is: a) ratio of height to weight of the contour of human subject b) quantized data of ratio c) distance between centroid points of upper and lower parts of the contour of human body, d) quantized data of c, e) recognized states/most likely decoded states (Viterbi paths), x axis is time (2nd subject's data).

In Fig.3, y axis represents: a) ratio of height to weight of the contour of the tracked human subject b) quantized data of ratio c) difference between centroid points of upper and lower parts of the contour of human body, d) quantized data of c, e) recognized states/most likely decoded states (Viterbi paths), x axis is time (2nd subject's data). From Fig.3 (e), we could see that recognized states are 2-1-3-1-2-4-1-4-1-5, which means almost all the states were correctly recognized. The correct rate of this case is 97.13%.

Table III shows the summarized results of the correct rate of human activity states of all subjects' data.

From Table III, we could see that states S1 and S5 (standing and lying down) were recognized by 100%, because human body shape information data of these states were stable and distinguishable from other states data. Walking state (S2) and sitting states (S3) were recognized by 97.0-99.0%. Miss recognition of slow walking is sometimes recognized by standing states, and also miss recognition of sitting is recognized by bending type 2. Another reason for miss recognition of states was occurred at transition from one state to another state. Overall correct rate of our human activity recognition process from Table III is 98.6-99.4%.

TABLE III. CORRECT RATE

Subject	Correct Rate S1	Correct Rate S2	Correct Rate S3	Correct Rate S4	Correct Rate S5	Overall
1	100.0	97.0	97.0	99.0	100	98.6
2	100.0	99.0	98.5	99.5	100	99.4
3	100.0	98.5	98	99.0	100	99.1
AVG	100.0	98.2	97.3	99.2	100	99.1
STDEV	0	1.04	0.76	0.28	0	0.40

It shows that our proposed human activity recognition process has high accuracy to represent right human activity, because it has mainly two reasons: first one is that we applied HMM, and another one is that our selection of several geometric parameters of human body is enough to represent human activities.

There have been many research works [7-16] of human activity and gait behavior recognition, and human identification through gait gesture using image processing techniques and attached some sensors to human. But they had some limitations, such as, offline recognition, cameras used are fixed at certain places, not dynamic, and so on.

Human activity recognition process applied on the expressible geometric parameters of the contour of the tracked human subject is the most important part of bio-monitoring mobile robot in home environment.

V. CONCLUDING REMARKS AND FUTURE WORK

We have developed algorithms for calculating several geometric parameters from the contour data of the tracked human subject extracted from the depth image of the Kinect sensor, vision for a mobile robot. Finally, we proposed a framework of human activity recognition from side viewpoint tracking by applying HMM based on preprocessed geometric parameters such as, the ratio of height to weight of the contour of the tracked human body and distance (pixels) between the centroid points of the upper and the lower parts of the contour data without any attached sensors and color markers to human body. This study is proposed human activity recognition by using the contour data of the tracked human subject instead of lower limb joints angle used in our previous study, that it helps not to use of the color markers attached to human lower limb joints. Applying HMM for the geometric parameters could achieve high accuracy, due to the fact that the contour data of the tracked human body expresses the complete information about human activities used in this study. Our proposed method brings the high rate of recognition of human activity and is effective in home environment and in real time.

In future, we will develop our own skeleton point extraction algorithm from the depth image.

REFERENCES

[1] Myagmarbayar Nergui, Yuki Yoshida, Wenwei Yu, (2012), "Human gait behavior interpretation by a mobile home healthcare robot", Journal of Mechanics in Medicine and Biology, World Scientific Publishing Company, Vol. 12, Issue 04, pp.1-24.

[2] Myagmarbayar Nergui, Nevrez Imamoglu, Yuki Yoshida, Wenwei Yu, (2012), "Human Gait Behavior Classification using HMM based on Lower Body Triangular Joint Features", The 14th IASTED International Conference on Signal and Image Processing, ACTA press, Honolulu, USA, pp. 212-219.

[3] H. Kwon, Y. Yoon, J. B. Park, et al, "Person tracking with a mobile robot using two uncalibrated independently moving cameras," in IEEE International Conference on robotics and Automation, Barcelona, Spain, 2005, pp.2877-2883.

[4] Meenakshi Gupta, Laxmidher Behera, et al, "A Novel Approach of Human Motion Tracking with the Mobile Robotic Platform," 13th International Conference on Modelling and Simulation, UKSim, 2011.

[5] Joao Cunha and Eurico Pedrosa, et al, "Using a Depth Camera for Indoor Robot Localization and Navigation," in Robotics Science and Systems (RSS) conference 2011, on Campus at the University of Southern California, poster presentation P5.

[6] Joydeep Biswas and Manuela Veloso, "Depth Camera based Localization and Navigation for Indoor Mobile Robots," www.cs.cmu.edu/~mmv/papers/11rssw-KinectLocalization.pdf, 2011.

[7] L.Wang, T.Tan, et al, "Silhouette Analysis-Based Gait Recognition for Human Identification," in IEEE transactions on pattern analysis and machine intelligence, 2003, Vol.25.

[8] Md. Z.Uddin, T.S.Kim et al, "Video-based Human Gait Recognition Using Depth Imaging and Hidden Markov Model: A Smart System for Smart Home," SHB2010-3rd International Symposium on Sustainable Healthy Buildings, Seoul, Korea, 2010.

[9] Naveen Rohila, et al, "Abnormal Gait Recognition," International Journal on Computer Science and Engineering, 2010, pp.1544-1551.

[10] M. Pushpa Rani1, et al, "An efficient gait recognition system for human identification using modified ICA," International Journal of computer science & information technology, 2010, Vol.2.

[11] Lawrence R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proceedings of the IEEE, 1989, Vol.77.

[12] Cuong Tran, Mohan Manubhai Trivedi, "3-D Posture and Gesture Recognition for Interactivity in Smart Spaces," IEEE Transactions on Industrial Informatics, 2012, Vol.8, No.1.

[13] Xianbin Cao, Bo Ning, Pingkun Yan,Xuelong Li, "Selecting Key Pose on Manifold for Pairwise Action Recognition," IEEE Transactions on Industrial Informatics, 2012, Vol.8, No.1.

[14] Ju Han, and Bir Bhanu, "Human Activity Recognition in Thermal Infrared Imagery" Computer Vision and Pattern Recognition - Workshops, CVPR Workshops. IEEE Computer Society Conference, June 2005

[15] Md. Atiqur Rahman Ahad, et al, "Human Activity Recognition: Various Paradigms" International Conference on Control, Automation and Systems, Seoul, Korea, Oct 2008.

[16] Md. Zia Uddin, Nguyen Duc Thang, Jeong Tai Kim, and Tae-Seong Kim, "Human Activity Recognition Using body Joint Angle Features and Hidden Markov Model", ETRI Journal, Aug 2011, Vol.33, No.4.