Multi-user Human Tracking Agent for the Smart Home

Juyeon Lee, Jonghwa Choi, Dongkyoo Shin, and Dongil Shin*

Department of Computer Science and Engineering, Sejong University, 98 Kunja-Dong, Kwangjin-Gu, Seoul, Korea
{jylee, jhchoi}@gce.sejong.ac.kr, {shindk, dshin}@sejong.ac.kr

Abstract. This paper presents a human tracking agent that recognizes the location and motion of the human in the home. We describe the architecture of the human tracking agent, and present an image recognition algorithm to track location and motion of the human. The human tracking agent decides the human's location, which changes in real-time, through the reletive distance of home furniture (or appliance) and human. Unlike the human's location, because a person's appearance (height, weight) is different for each person, a human's motion should be recognized to be different from each other person. We converted the image (that is acquired from the network camera) into a standard image (that is defined in the human tracking agent) for recognition of multi-user's motion. We used a LSVM(linear support vector machine) to recognize the feature patterns for human motion. In our experiment, results of motion recognition showed excellent performance accuracy of over 80%.

Keywords: User Tracking System, Smart Home, User Motion Recognition.

1 Introduction

A home network that integrates sensors, actuators, wireless networks and context-aware middleware will soon become part of our daily life [1]. We define this environment as a smart home [2]. A smart home is a house or living environment that contains the technology to allow devices and systems to be controlled automatically. Also, one of the goals of a smart home is to support and enhance the abilities of its occupants to execute tasks. An artificial intelligence agent in a smart home learns about the occupants and the smart environment, and predicts the appliance services that they will want. That is, it means an intelligent space that provides a suitable service that is predicted from various home contexts that happen in the ubiquitous smart home. The most important contexts among all contexts, which happen in home, are the human's location and motion. Many researchers used various sensor devices (IrDA, ultra audible sound, radio communication signal and the smart floor that analyzes user's location through pressure sensors) [3-9]. But, the various sensors that are presented have the shortcoming that they must attach a communication tag on the user's body. A location recognition method that uses a camera does not attach a communication tag on the human's body, and recognizes the human's motion as well as the human's location from the image offered by the camera. The location recognition system that uses IrDA establishes IrDA sensors in the office, and recognizes the user's location through a signal communication active badge that is

* Corresponding author.

Z. Shi and R. Sadananda (Eds.): PRIMA 2006, LNAI 4088, pp. 502 – 507, 2006.

attached to the user [3]. Each active badge has a unique serial number and transmits a serial number to the IrDA sensor that is attached in the office. A location recognition system that uses ultrasonic audible sound has been studied by the University of Cambridge (Active Bat) and MIT (cricket system) [4, 5]. The Active Bat attaches an ultrasonic generator (which is called the Bat) to a person or objects, and attaches an ultrasonic receiver in the office. The cricket that is developed by MIT is a location recognition system based handset. That is, the ultrasonic generator is attached in the office, and the person has an ultrasonic receiver. There is a system named the RADAR developed by Microsoft that is a representative method of the use of RF signals [6]. The Aware Home establishes pressure sensors in the floor, and recognizes user location [7]. The MavHome implemented a user's routing algorithm through use of the pressure sensors in the floor [8, 9].

2 Structure of Human Tracking Agent

Figure 1 shows the structure of our home tracking agent.

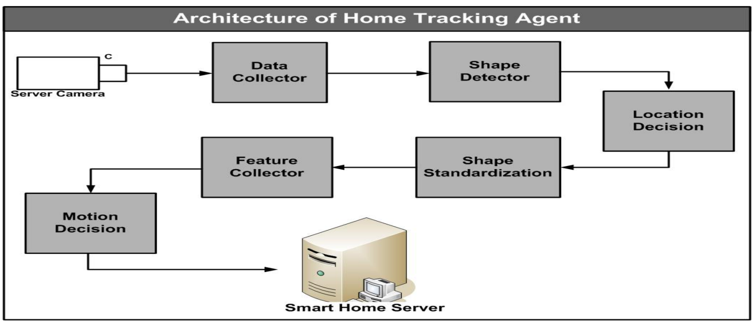

Fig. 1. Architecture of the home tracking agent

The Data Collector module acquires color images (720 X 486) from the digital network cameras every three seconds. The Shape Detector module analyzes the human's volume from four images, which is acquired from the Data Collector module. The Shape Detector module calculates the user's absolute coordinates in home and relative coordinates with reference to the furniture and home appliances. We used a moving window to extract the human's coordinate in each image. The Location Decision module analyses the user's location through coordinates that are offered by the Shape Detector module. The Location Decision module analyses whether the human is near to which home appliance (or furniture) through comparison between the absolute coordinate of human and absolute coordinate of appliance (or furniture) in the home. The Shape Standardization module converts the image (that is acquired by the network camera) into a standard image (that is defined in the human tracking agent) for recognition of multi-user's motion. The Feature Collector module extractes the image that is changed by the Shape Standardization module, and the Motion Decistion module predicts the motion of the human.

We applied three kinds of images to acquire absolute coordinates and relative coordinates in the home. The first image is an image in which the user and furniture (and home appliances) are excluded. The second image is an image in which furniture and home appliances are arranged in the ubiquitous smart home and the third image is an image that includes the human inside the second image. To calculate the human's absolute coordinates in the home, we apply subtraction of three images, and determine the x and y coordinates of the moving window that include the human in the image. When the home tracking agent starts, it executes a subtraction between the first image (which is excluding human, furniture and home appliances) and the second image (which is including furniture or home appliances), and calculates the absolute coordinates of furniture and home appliances using the moving window. The Shape Detector module receives raw image from the Data Collector module, and distinguishes the human's image through subtraction between the second image and third image (that is, including the human with second image), and then calculates the human's absolute coordinates in the home and human's relative coordinates with reference to the furniture and home appliances. It also decides whether a human is near to which furniture (or home appliance). If there is a human in an important place that is defined to the human tracking agent, the Shape Detector module calculates relative coordinates of the moving window, which includes a human, and transmits those to the Location Decision module. We defined the significant places in the home which include sofa, desk, bed, etc. The Location Recognition module can judge a multi-user's location without conversion of the acquired image. Unlike the human's location, because person's appearance (height, weight) differs for an each person, the human's motion should be recognized to differ in case of each person. The Shape Standardization module takes charge of motion recogntion in the multi-user situation. To recognize the human's motion, we defined a feature sets of six standardized motion. Figure 2 shows six motions that is recognized in the human tracking.

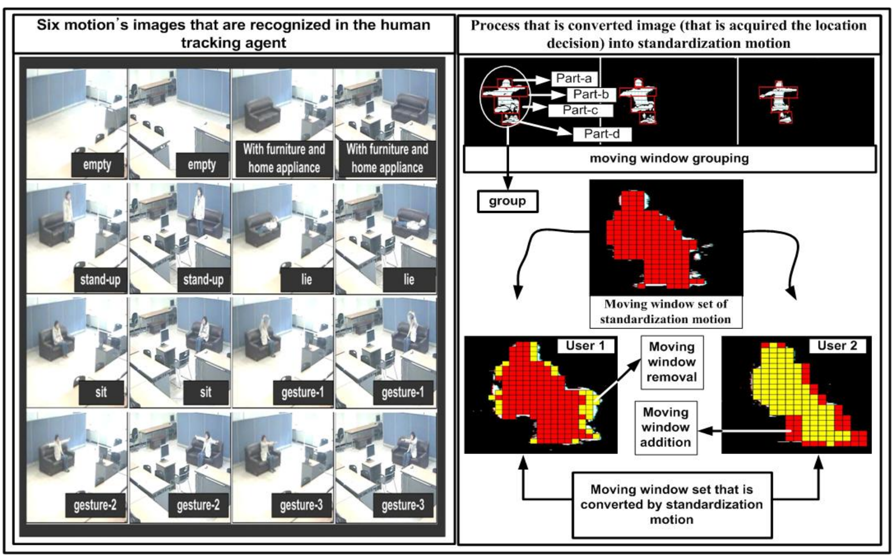

Fig. 2. Six motions that are recognized in the human tracking agent and the process that converts the image (that is acquired the location decision) into standardized motion

The human tracking agent predicts six human's motions (stand-up, lie, sit, gesture-1, gesture-2 and gesture-3). The Shape Standardization module groups moving window that is acquired from the Location Decision module. The algorithm is as follows. The first, count each row's column number and group rows that total column's sum is same (error range is +/- 2). Second, add or reduce moving window of each group by comparison between standard motion and motion that is acquired. Third, if count of moving window that is removed (or added) is an even number, the Shape Standardization module removes (or adds) first and last moving window (column) repeatedly, otherwise, removes (or adds) last moving window one more.

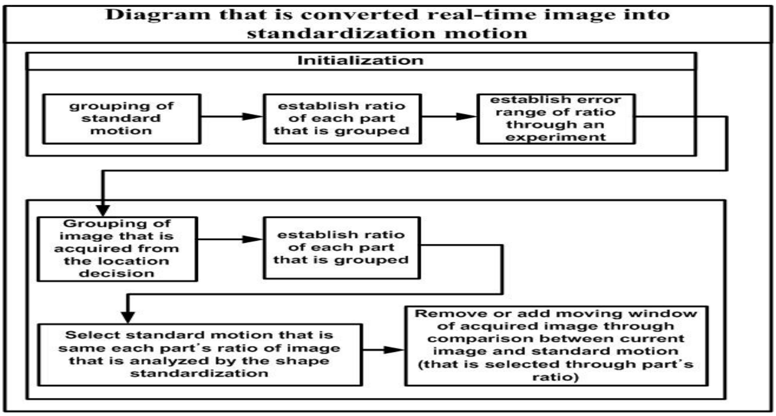

Fig. 3. Diagram that convert real-time image into a standardized motion.

Figure 3 shows diagram that converts real-time image into standardized motion. The feature set that is transmitted from the Shape Standardization module is normalized between 0.1 and 0.9 in the feature collector. And, the Motion Decision module applied normalized features as input values for the LSVM. The Motion Decision module that is presented recognizes the first step motion (sit, lie, stand-up) and second step motion (three gestures).

3 Implementation and Evaluations

Our laboratory is 11 m long, 12 m wide and 3 m high. We use network digital cameras to acquire the human's image. The network cameras that are installed in the laboratory include error data, and we overcame this problem through extension of data scope. Nevertheless, the error range of the human's location recognition is 0.024 m. Figure 4 shows subtraction images and the human motion pictures that are acquired in the laboratory.

We experimented with the performance of motion recognition in the human tracking agent, and evaluated performance of the human tracking agent through comparison between single-user and multi-user conditions. Table 1 shows the

Fig. 4. Pictures of human motion that are acquired in the laboratory

	Number of Support vector	Number of kernel evaluations	Norm of longest vector	Precision on test(%)
Sit	68	13542	2.24578	89
Stand	71	13477	2.32457	91
Lie	64	13356	2.28451	92
Sit_gesture1	65	13487	2.34622	88
Sit_gesture2	68	13587	2.38457	89
Sit_gesture3	69	13257	2.19874	82

Fig. 5. Experimental results with the human tracking agent that recognizes only a single-user

	Number of Support vector	Number of kernel evaluations	Norm of longest vector	Precision on test(%)
Sit	57	12992	2.21434	81
Stand	65	13023	2.28132	84
Lie	54	13112	2.27878	85
Sit_gesture1	59	12843	2.31436	79
Sit_gesture2	61	12934	2.39105	80
Sit_gesture3	61	13998	2.21769	76

Fig. 6. Experimental results for the human tracking agent that recognizes multiple users

performance of motion recognition by the human tracking agent that is limited to single-user operation. The six kinds of motion that are recognized in the human tracking agent are shown in Figure 5 to have a high average accuracy rate of 88.5%. The sit_gesture3 shows a lower accuracy rate of 82% because the feature set of the sit_gesture3 does not have a feature that is clearly distinguishable from other motions.

Figure 6 shows the performance of the human tracking agent that recognized multiple users. The human tracking agent decreases performance of motion recognition in multi-user mode in comparison with recognition in single-user mode. But, still, results of motion recognition showed an excellent performance of over 80%.

We are currently studying effective motion recognition system by using the feature's relative importance.

4 Conclusions

This paper presents a human tracking agent that recognizes the location and motion of the human in the home. The human tracking agent decides the human's location, which changes in real-time, through the reletive distance of home furniture (or appliance) and human. Unlike the human's location, because a person's appearance (height, weight) is different for each person, a human's motion should be recognized to be different from each other person. We converted the image (that is acquired from the network camera) into a standard image (that is defined in the human tracking agent) for recognition of multi-user's motion. The human tracking agent that is presented in this paper support multi-user motion recognization. We used a LSVM to recognize the feature patterns for human motion. In our experiment, results of motion recognition showed excellent performance accuracy of over 80%.

References

1. Schulzrinne. H, Xiaotao Wu, Sidiroglou. S, Berger. S.: Ubiquitous computing in home networks. Communication Magazine. IEEE, Vol. 41, Issue. 11, (2003) 128-135
2. Sherif, M. H.: Intelligent homes: a new challenge in telecommunications standardization. Communication Magazine. IEEE, Vol. 40, Issue.1, (2002) 8-8
3. Want, R., Hopper, A.: Active badges and personal interactive computing objects. Consumer Electronics. IEEE Transactions on Vol. 38, Issue. 1, (1992) 10 - 20
4. Hazas, M.; Hopper, A.: Broadband Ultrasonic Location Systems for Improved Indoor Positioning. Mobile Computing. IEEE Transactions on Vol. 5, Issue. 5, (2006) 536 - 547
5. Adam Smith, Hari Balakrishnan, Michel Goraczko, Nissanka Bodhi Priyantha.: Tracking Moving Devices with the Cricket Location System. 2nd International Conference on Mobile Systems. Applications and Services (Mobisys 2004), (2004)
6. p.Bahl and V.Padmanabhan.: RADAR: An In-Building RF-based User Location and tracking system. Proc, IEEE infocom, IEEE CS Press, Los Alamitos, CA, (2000) 775-784
7. Orr, J. Robert, D. Gregory, Abowd.: The smart floor: A Mechanism for natural user identification and tracking. To appear in the proceedings of the 2000 Conference on Human Factors in Computing systems (CHI 2000), The Hague, Netherlands, (2000)
8. MavHome, http://cygnus.uta.edu/mavhome/
9. Das. S.K, Cook, D.J, Battacharya. A, Heierman. E.O. III, Tze-Yun Lin.: The role of prediction algorithms in the MavHome smart home architecture. Wireless Communications. IEEE, Vol. 9, Issue. 6, (2002) 77 - 84