Toward a Companion Agent for the Elderly – The Methods to Estimate At-Home and Outside-Home Daily Life Activities of the Elderly Who Live Alone

Yuma Takeda, Hung-Hsuan Huang*, and Kyoji Kawagoe

Graduate School of Information Science & Engineering, Ritsumeikan University, Japan
hhhuang@acm.org

Abstract. With advances in medical technology, people's life have been extended, and there are more and more older adults isolated. If they do not maintain social life with others, they may feel loneliness and anxiety. For their mental health, it is reported effective to keep their social relationship with others, for example, the conversation with their caregivers or other elderly people. Active listening is a communication technique that the volunteer listener listens to the speaker (the elderly) carefully and attentively by confirming or asking for more details about what they heard. This helps to make the elderly feel cared and to relieve their anxiety and loneliness. This paper presents our in-progress project aiming to develop a framework of a virtual companion agent who is always with the user and can engage active listening to maintain a long-term relationship with elderly users. In order to achieve the agent's companionship with the user for a longer period, we believe that it is essential to make the agent to understand the user as best as it can. This kind of user-fitted conversation is not addressed in previous companion agent work. The proposed approach is the acquisition of the "memory" of the user's daily life in two situations, at-home and outside-home. In the former one, multiple Microsoft Kinect depth sensors were adopted. The depth information is integrated to detect the user's position and posture and then to estimate the user's daily activity. In the outside-home configuration, the prototype application is an Android smartphone application that recognizes the user's moving status with the information from the on-board three-axis accelerometer as well as the location of the user from GPS information. These data are then used to estimate the user's outside-home activity. All estimated daily activities are recorded in an activity history database. Both the at-home and outside-home activity estimation methods have been developed and have been evaluated in a laboratory environment with student subjects at a moderate accuracy. The interface of the companion agent is being designed with the results from human-human and human-agent (driven by the data from the human listener condition) subject experiments. After the technologies are more matured, we would like to conduct real-world experiment with elderly subjects in near future.

1 Introduction

With advances in medical technology, the average life expectancy of world population is increasing. Since the probability of becoming cognitively impaired increases with age

* Corresponding author.

C. Stephanidis and M. Antona (Eds.): UAHCI/HCII 2014, Part III, LNCS 8515, pp. 395–402, 2014.

(roughly 10% of over 65 years old people), one side effect of increasing life expectancy is the emerging number of dementia patients. It is said that currently there are already around 40 million dementia patients all over the world. This is particularly severe in developed countries where the problem of aging population proceeds. Japan, probably is the country in most severe situation in the world. According to a recent statistical data, the number of dementia patients in Japan has already exceeded two million (1.6% of the population), and the number will keep increasing to 4.5 million (4.1% of the population) by 2035.

If they do not maintain social life with others, they may feel loneliness and anxiety. For their mental health, it is reported effective to keep their social relationship with others, for example, the conversation with their caregivers or other elderly people. Reminiscence or life review [1] is a well known method to slow the progress of the most prominent symptom of dementia, memory impairment. It is also reported in the literature [2,3] that repetitive stimuli on cognitive functions in the environment is also effective in suppressing the degradation of specific cognitive abilities.

Active listening is a communication technique that the listener listens to the speaker carefully and attentively by confirming or asking for more details about what they heard. This kind of support helps to make the elderly feel cared and to relieve their anxiety and loneliness. However, due to the lack of the number of volunteers comparing to that of the elderly who are living alone, the volunteers may not be always available when they are needed. In order to improve the effect, always-available and trustable conversational partners in enough number are demanded. This paper presents our in-progress project aiming to develop a virtual companion agent who can engage active listening and maintain a long-term relationship with elderly users.

This paper presents our approaches in acquiring the "memory" of the user's daily life in two situations, at-home and outside home. Because of the different level of constraints of the situations, we tried to maximize the richness of sensory information with different sensor technologies for each of them. Fig. 1 shows the conceptual diagram of this project, where the companion / listener agent can utilize the daily activities of the user and engage the conversation with him / her. The daily activity database are created from the information gathered by portable device (smartphone) and at-home sensors (Microsoft Kinect). These information can be further searched by medical institutions or the family of the users from remote.

2 Related Work

Various assistive technologies for dementia patients have been proposed so far. Since it is difficult to find a sufficient number of caregivers for dementia patients in many countries, besides providing physical assistance to those with physical impairments, it is important for assistive artifacts to provide communication functions [4]. An embodied conversational agent can effectively serve as a listener for people with dementia if it is accepted as a companion by the patients. Previous studies on the acceptance of such an agent by elderly people reported that it is important for the agent to display social signals, like smiling and head nods [5]; this enables the agent to gain the patient's trust and enhances intimacy [6].

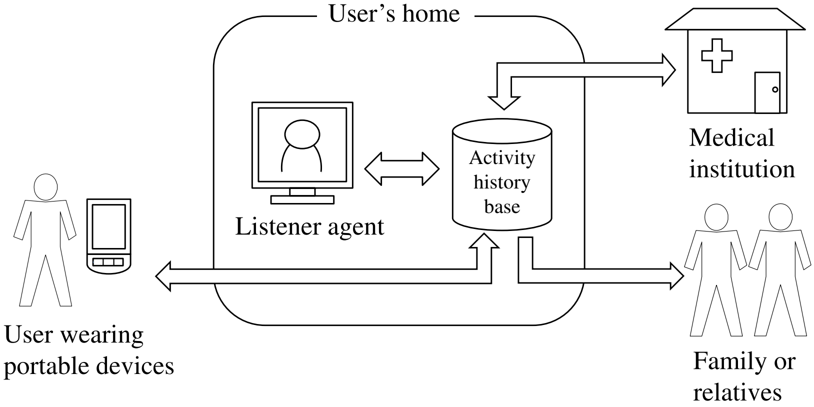

Fig. 1. Conceptual diagram of the project

In the aspect of conversational companion agents, most previous studies focused on user perception of empathy and affectiveness during the interaction with the agent. Kanoh et al. [7] investigated user acceptance of a robot in recreational use inside health care facilities for elderly people. Although the participants showed positive reactions to the robot, the interaction between the participants and the robot was seldom observed. Bickmore et al. [4] investigated the effects of verbal and non-verbal empathic behaviors of a 2D graphic agent and found that the subjects did rate the agent more caring if it shows those behaviors. Leite et al. [8] investigated a robot cat showing empathic behaviors (voice, facial expressions, and head movements) on the players of chess game. Smith et al. [9] proposed the integration of affective dialogue with a deliberative architecture. These studies showed that the display of empathic behaviors can usually make the conversational artifacts better accepted by users, which is a requirement of artificial companions. However, these proposed systems are neither equipped with the mechanism to keep long-term relationship with the users nor used in serious tasks.

Other studies try to model episodic memory which is essential to maintain the dialogue with users in long-term relationship. Sieber and Krenn [14] proposed a W3C RDF (resource description framework) based presentation of past interaction and user preferences. In order to achieve higher efficiency and more realistic dialogue, Lim et al. [10] integrated "forgetting" feature into their episodic memory model. Campos and Paiva [11] proposed a chat agent for assisting a teenager user on self-reflection about what happens in his / her life. The dialogue is pro-active and adapted to the main goals of a teenager user, school, love and play. These projects also do not aim serious use of companion agent, but the Campos' work shares similar general idea with us, i.e. store and acquire personal memory by the interaction with a companion agent.

3 Recording of Daily Activities

In order to achieve the agent's companionship with the user for a longer period from several months to several years, we believe that it is essential to make the agent to understand the user as best as it can. Previous approaches include gathering the profile information of the user in advance and record the interaction history between the agent and the user. In addition to them, tracking the activity history and the events occurred in the user's daily life seems to be a reasonable approach if the agent is always with the user. By tracking the regular patterns as well as the occurrences of irregular patterns, the agent may discover the user's personality and habits and have more chances to engage the conversation with the user. Then the agent can then trigger the utterances like "You waked up latter today. Did you feel bad somewhere?" or "Please take care of yourself better" if it finds that the user eats out everyday.

3.1 At-Home Situation

In the at-home configuration, Microsoft Kinect depth sensors were adopted because the balance between its effectiveness and cost, as well as the user can be free of attaching some dedicated sensors on his/her body. Due to the fact that a single Kinect can only detect the distance between itself and the objects within a range between 0.8 and 4 m, multiple Kinects are required to cover a typical one-room apartment (about 30 m^2) in Japan. The method to integrate the coordinate systems of two Kinects to the world coordinate system is evaluated to have a precision with errors less than 0.6 m in a simulated room. This should be enough to detect the locations of the user inside his/her home. From the location and prior knowledge of the room layout (the locations of TV, toilet, kitchen, etc), we expect that we can estimate the user's at-home activities. Fig. 2 shows the layout of the simulated room where we conducted the experiments. We measured the precision of the position estimation method at the preventative positions with a two-Kinect setup. Table 1 shows the results. The precision varies while the distance toward Kinects. The precision is only at moderate level but should be enough to distinguish the spaces where the user doing his / her activities.

Table 1. Measured precision of the position estimation method for at-home situation

ID	(X, Y)	estimated (X, Y)	error
1	(-2.000, -0.500)	(-1.121,-0.278)	1.136
2	(-0.500, -1.000)	(-0.331, -0.559)	0.471
3	(-1.000, -0.500)	(-0.580, -0.235)	0.496
4	(0.000, 0,000)	(0.591, -0.170)	0.524
5	(-0.500, 0.500)	(0.057, 0.084)	0.102
6	(-0.500, 0.500)	(-0.304, 0.323)	0.263
7	(0.500, 0.500)	(0.284, 0.333)	0.271
8	(-1.000, 1.000)	(-0.551, 0.473)	0.691
9	(1.500, 1.000)	(0.001, 0.004)	1.799
10	(0.000, 1.500)	(-0.024, 0.900)	0.600

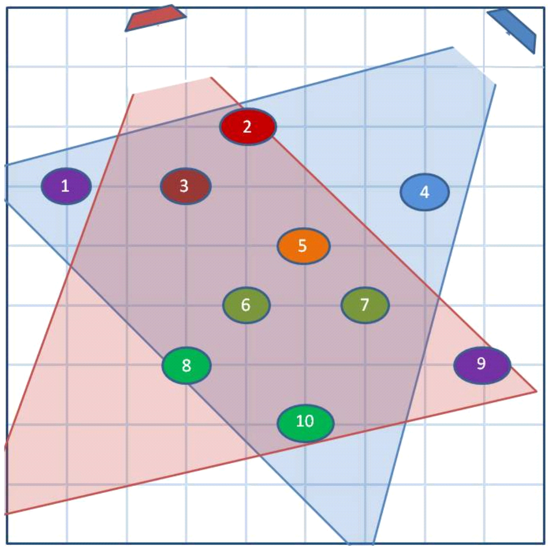

Fig. 2. Coordinations where the precision of position estimation was measured

3.2 Outside-Home Situation

In the outside home configuration, the prototype application is implemented on a Samsung Galaxy SII phone. This Android application recognizes the user's moving status (walking, running, bicycle, car, or train) with the information from the on-board three-axes accelerometer (sampling rate: 100 Hz). The recognition process uses a priorly trained C4.5 decision tree by Weka data mining tool [12] from 30-minute training data of each class. Since there should be no difference of feature values among different person on transport vehicles, those training data were collected from one person. On the other hand, the walking data were collected from five college students (three males and two females). The features used are the maximum, average, and deviation of each axis. The measurement is based on a 5-second sliding window in real-time, and a 92.2% 10-fold cross validation accuracy is achieved. Base on this mechanism, we further measure the activities in larger temporal granularity, i.e. 10 minutes, one hour, and one day. The preliminary experiment is done with one male college student's activities in one month. In addition to the features for detecting moving status, other features like time period and the types of facility where the subject is in were used. Table 2, 3, and 4 show the confusion matrices of the classification results of each granularity, respectively. The classification accuracy at 10-fold cross validation was shown in Table 5. Furthermore, these data were sent to a database where queries of the user activities is possible from Web interface (Fig. 3).

The application also logs the user's current position from the location information of on-board GPS sensor (sampling rate: 1 Hz). The moving status and position is sent to a back end server in trunks periodically (for example, once every 10 minutes). Our next

Table 2. Confusion matrix of 10-minutes activity estimation. The data in columns are the classification results

	Lunch	Dinner	Desk work	Restaurant	Shopping	Walking	Bicycle	Car	Train
Lunch	124	0	5	1	1	0	0	0	0
Dinner	0	127	5	0	1	0	0	0	0
Desk work	11	33	118	0	0	0	0	0	0
Restaurant	0	0	0	174	0	0	0	0	0
Shopping	0	0	0	0	173	0	0	0	0
Walking	0	0	0	0	0	150	0	0	0
Bicycle	0	0	0	0	0	0	156	0	0
Car	0	0	0	0	0	0	0	135	0
Train	0	0	0	0	0	0	0	0	138

Table 3. Confusion matrix of one-hour activity estimation. The data in columns are the classification results

	Meal	Shopping	Desk work	Moving
Meal	76	1	19	0
Shopping	16	91	5	0
Desk work	19	0	83	0
Moving	0	0	0	100

Table 4. Confusion matrix of one-day activity estimation. The data in columns are the classification results

	Study & Research	Meal	Recreation
Study & Research	10	0	0
Dinner with friends	0	8	2
Recreation	0	2	8

Table 5. Classification accuracy (10-fold cross validation) in different temporal granularities

Time Slice	Moving	10 minutes	One hour	One day
Accuracy	92.2%	95.9%	85.3%	86.6%

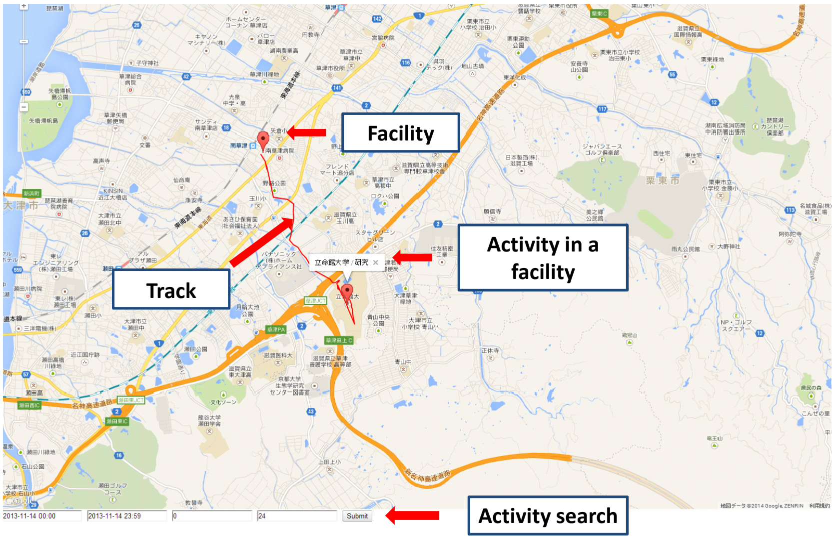

Fig. 3. Visualization of the activities of outside-home situation

step is to extend the companion agent interface to the smartphone. Considering safety issue, it is not necessary to include the graphical character. However, as the literature reports that the user can feel the agent migrate even its form changes [13], it would be easier to establish trustworthy relationship with the agent if the user can feel the agent is still with the user when (s)he is outside home. For example, using the same voice and the same personality model in both the at-home agent kiosk and the mobile phone.

4 Conclusions and Future Work

This paper presents a part of work of an ongoing project that aims to develop a virtual companion agent for the elderly. It is believed that the tracking and utilizing the daily life activity of the user for the agent's action decision-making can help to develop long-term relationship with the user. Kinect depth sensors were used in at-home situation while Android mobile phone is used in outside home situations to record the user's locations. These data are further used to estimate the user's activities. As future work, first of all, we would like to integrate the activity information from both the at-home and outside home situations and develop an uniformed memory representation for the agent. After that, we would like to complete the development of the interaction loop of user activity recognition and agent behavior generation. Finally, we would like to deploy the complete system in long-term practical use to evaluate its effectiveness.

References

1. Butler, R.N.: Successful aging and the role of the life review. Journal of the American Geriatrics Society 22(12), 529–535 (1974)
2. Lazarov, O., Robinson, J., Tang, Y.P., Hairston, I.S., Korade-Mirnics, Z., Lee, V.M.Y., Hersh, L.B., Sapolsky, R.M., Mirnics, K., Sisodia, S.S.: Environmental enrichment reduces Aβ levels and amyloid deposition in transgenic mice. Cell 120(5), 701–713 (2005)
3. Kempermann, G., Gast, D., Gage, F.H.: Neuroplasticity in old age: Sustained fivefold induction of hippocampal neurogenesis by long-term environmental enrichment. Annals of Neurology 52(2), 135–143 (2002)
4. Bickmore, T.W., Picard, R.W.: Towards caring machines. In: Proceeding CHI EA 2004 CHI 2004 Extended Abstracts on Human Factors in Computing Systems, pp. 1489–1492 (April 2004)
5. Heerink, M., Kröse, B., Evers, V., Wielinga, B.: Studying the acceptance of a robotic agent by elderly users. International Journal of ARM 7(3), 33–43 (2006)
6. Cassell, J.: Nudge Nudge Wink Wink: Elements of Face-to-Face Conversation for Embodied Conversational Agents. In: Embodied Conversational Agents, pp. 1–27. The MIT Press (2000)
7. Kanoh, M., Oida, Y., Nomura, Y., Araki, A., Konagaya, Y., Ihara, K., Shimizu, T., Kimura, K.: Examination of practicability of communication robot-assisted activity program for elderly people. Journal of Robotics and Mechatronics 23(1), 3–12 (2011)
8. Leite, I., Mascarenhas, S., Pereira, A., Martinho, C., Prada, R., Paiva, A.: "Why can't we be friends?" an empathic game companion for long-term interaction. In: Safonova, A. (ed.) IVA 2010. LNCS, vol. 6356, pp. 315–321. Springer, Heidelberg (2010)
9. Smith, C., Crook, N., Boye, J., Charlton, D., Dobnik, S., Pizzi, D., Cavazza, M., Pulman, S., de la Camara, R.S., Turunen, M.: Interaction strategies for an affective conversational agent. In: Safonova, A. (ed.) IVA 2010. LNCS, vol. 6356, pp. 301–314. Springer, Heidelberg (2010)
10. Lim, M.Y., Aylett, R., Ho, W.C., Enz, S., Vargas, P.: A socially-aware memory for companion agents. In: Ruttkay, Z., Kipp, M., Nijholt, A., Vilhjálmsson, H.H. (eds.) IVA 2009. LNCS, vol. 5773, pp. 20–26. Springer, Heidelberg (2009)
11. Campos, J., Paiva, A.: May: My memories are yours. In: Safonova, A. (ed.) IVA 2010. LNCS, vol. 6356, pp. 406–412. Springer, Heidelberg (2010)
12. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: An update. ACM SIGKDD Explorations 11(1), 11–18 (2009)
13. Ogawa, K., Ono, T.: ITACO: Effects to interactions by relationships between humans and artifacts. In: Prendinger, H., Lester, J.C., Ishizuka, M. (eds.) IVA 2008. LNCS (LNAI), vol. 5208, pp. 296–307. Springer, Heidelberg (2008)
14. Sieber, G., Krenn, B.: Towards an episodic memory for companion dialogue. In: Safonova, A. (ed.) IVA 2010. LNCS, vol. 6356, pp. 322–328. Springer, Heidelberg (2010)