A Support System for Context Awareness in a Group Home Using Sound Cues

H. Kanai[1], T. Nakada[2], Y. Hanba[3], S. Kunifuji[2]
[1]Center of Knowledge Science, Japan Advanced Institute of Science and Technology, Ishikawa, Japan
[2]School of Knowledge Science, Japan Advanced Institute of Science and Technology, Ishikawa, Japan
[3]Solution Division, Ricoh Software Inc., Tokyo, Japan

Summary

Objectives: In a group home, caregivers should be aware of the inhabitant's real-time situation. The aim of our study is to facilitate the awareness of an inhabitant's situation by means of enhanced sound cues.

Methods: We propose an audio notification system that indicates the real-time situation of persons in a group home environment using sound cues instead of visual surveillance. The notification system comprises a prediction and a notification function. The prediction function estimates a person's real-time situation using a Bayesian network and sensed information; the notification function informs recipients of the predicted situation and the confidence level of the prediction by means of sound cues. We use natural sounds as sound cues.

Results: As a first step to examine our system in a group home, we conducted operation and performance tests of each unit under a simple test environment. The correct prediction of the subject's situation is approximately 90%; further, it is shown that the sound cues should be selected according to their environmental dependence.

Conclusions: The results show that the method is useful for monitoring persons. As future study, we will conduct a field test on an implemented system and improve it for practical use in a group home.

Keywords

Group home, context awareness, ambient notification system, support for caregivers

Methods Inf Med 2008; 47: 198–202

doi:10.3414/ME9110

1. Introduction

Currently, the population of Japan is rapidly aging [1]. One of the measures employed in Japan to mitigate this situation is the establishment of many group homes for people with dementia. A group home is a type of home-based care service for elderly citizens with dementia. It is a facility where elderly citizens with dementia reside together and receive nursing care services such as meals and baths on a 24-hour basis. One caregiver is responsible for six or seven inhabitants, and a few caregivers share the caretaking duties of a house around the clock. The caretaking of the inhabitants should be performed on an individual basis. Hence, it is important for caregivers to be aware of an inhabitant's real-time situation in a group home. Many caregivers attempt to understand an inhabitant's situation to some extent by listening to some of the sounds emanating from his/her location. This use of sound enables the caregivers to understand the inhabitant's situation while carrying out other caretaking activities. However, inexperienced caregivers and those who are new to a group home find it difficult to perceive sounds that emanate from an inhabitant's location while they are engaged in other caretaking activities.

Related work on using sound for event notification includes the weakly intrusive ambient soundscape (WISP) system. It uses natural sounds and represents the sound intensities by changes in the playback volume and level of reverb [2]. This approach when applied in our case might cause recipients to inaccurately judge a situation. Our system introduces natural sounds with an acoustic noise (hissing noise) because it can maintain the sound volume without any processing effects. In another paper, to focus attention on a specific user, personalized music with altered instrumental sounds and rhythm is used [3]. Our system does not use these personalized sounds because it is intended to notify the situation of inhabitants to several caregivers. In a third paper, a system provides musical cues corresponding to the individual's affective state that is estimated by his/her facial movements [4]. Our system focuses on predicting an individual's situation based on his/her actions.

In this paper, we describe a notification system that indicates the real-time situation of people in a group home environment by means of sound cues. This system comprises a prediction and a notification function. The prediction function estimates a person's real-time situation by using a Bayesian network and sensed information; the notification function provides information on the predicted situation and the confidence level of the prediction by means of sound cues.

The following section describes our notification system, particularly a method that predicts a situation and notifies it to the user. The next section describes a user test of the implemented system in a simple test environment.

2. System

The structure of our system is shown in Figure 1. It comprises a sensing unit, situation prediction unit, and situation notification unit. The sensing unit detects the state of a person and objects in real time. The situation prediction unit estimates the person's situation by using a Bayesian network [5] along with the sensed information. The situation notification unit informs the users of the predicted situation by means of sound cues.

2.1 Situation Prediction Unit

We use a Bayesian network to predict a person's situation. The Bayesian network is a well-known model employed to represent uncertainty. It is a type of probabilistic graphical model. The network represents the joint probability distribution of a set of variables with explicit independence assumptions. In our study, the Bayesian network depends on the behavior of persons in a living environment. In order to predict a person's situation by using the Bayesian network, it is required that some parameters have a causal correlation with the person's situation. The nodes "Location", "Time", and "Status of Objects" correspond to these parameters; Location denotes the position of the person in a room, Time denotes the instant when the sensor detects the person's action and an object's movement, and Status of Object denotes the status of the objects in the rooms. As an example of "Location", "take a bath" is linked with "a bathroom". As an example of "Time", "performing ambulatory automatism" is linked with "midnight" in the case of individuals with dementia. As an example of "Status of Object", "a person is getting hungry or shows interest in a meal" is linked with "opening and closing motion of a refrigerator and a store cupboard".

When the states of these nodes are reflected in the Bayesian network, the confidence level of each person's predicted situation is calculated. We assume that the state corresponding to the highest confidence level denotes the real-time situation of the target. Figure 2 shows a Bayesian network for "Sleep". Here, the arrows represent the causal relationship between the "Time" and "Situation" nodes, and that between the "Situation" and "Location" nodes. The node Sleep is linked with three nodes, namely, Time, Location: Bed, and Location: Dining table. This implies that the confidence level of Sleep depends on the states of these nodes. While updating the states of these nodes in the Bayesian network, the confidence level of each node is transmitted in both directions. For example, if the Time is "midnight", the confidence level of Sleep will increase. Moreover, if the target individual is in bed, the confidence level will be further increased.

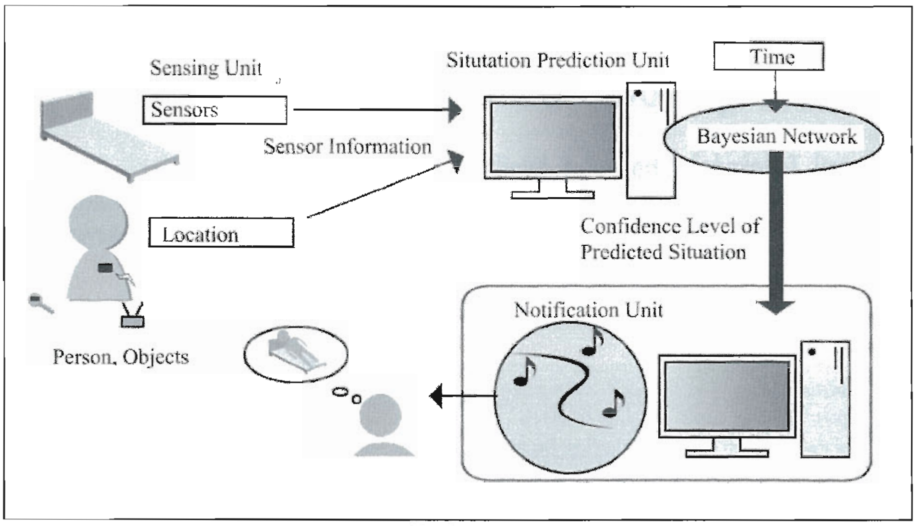

Fig. 1 Structure of our system

2.2 Situation Notification Unit

The situation notification unit translates the predicted situation of a person and its confidence level into sound cues. The unit provides context awareness regarding the person to recipients. In our study, natural sounds with an acoustic noise are used as sound cues. Here, natural sounds imply those sounds that originate from original or simulated actions and events. The natural sounds have a close relation to events occurring in the real world, for example, the sound made by an intercom when a person visits, typing sounds, ring alerts, and so on. We consider that these natural sounds facilitate the recognition of events by people. An acoustic noise such as a hissing noise represents the confidence level of the predicted situation. That is, if the confidence level of the predicted situation is high, the volume of the acoustic noise will be low. As a result, recipients will be able to clearly perceive the predicted situation and its confidence level.

The reasons for introducing the abovementioned sound cues are to maintain the

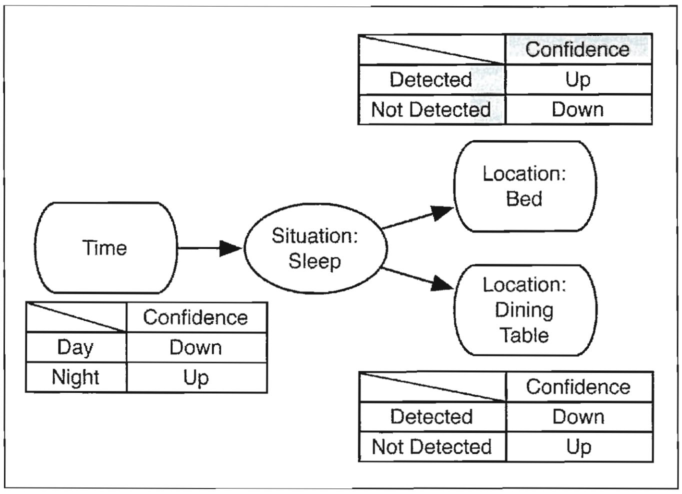

Fig. 2 An example of the Bayesian network model for "Sleep"

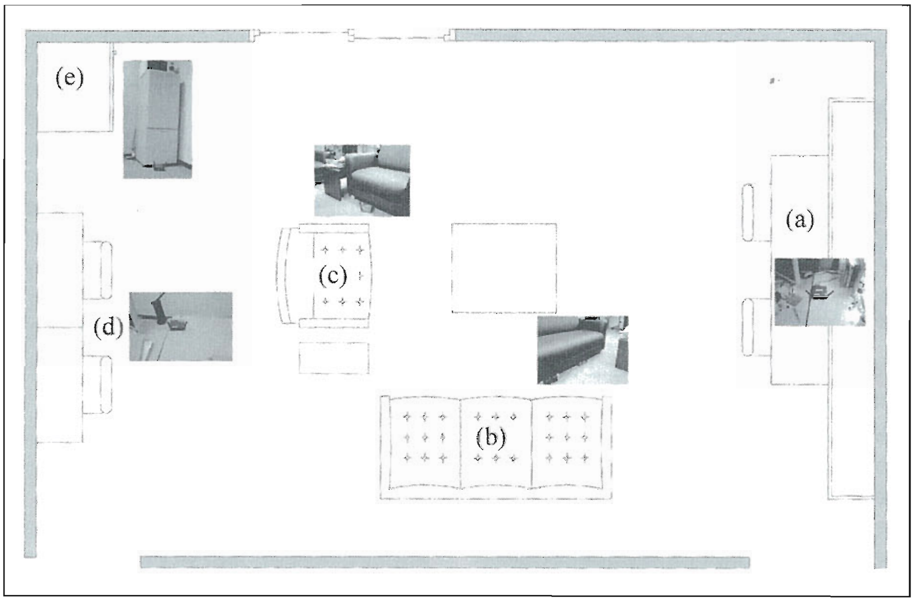

Fig. 3 Simple test environment

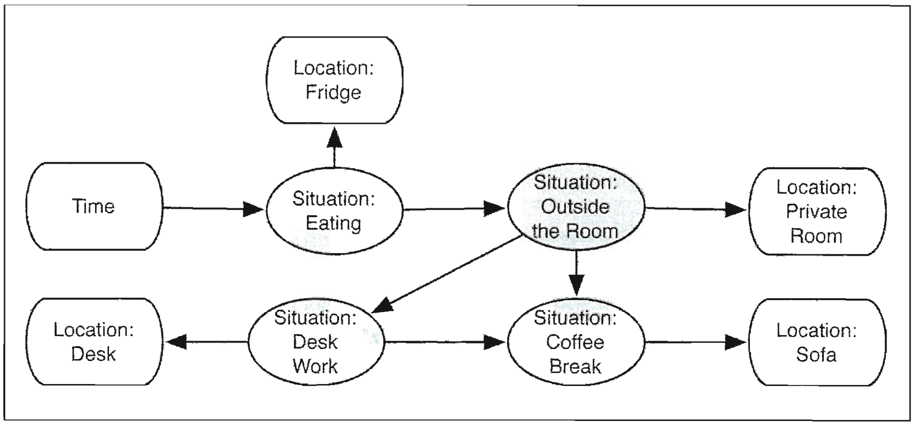

Fig. 4 Bayesian network model used in this investigation

volume of natural sounds at a single level and to maintain acoustic quality. If the predicted situation is represented by an alteration in the volume of natural sounds, a low-volume sound might be suppressed by other sounds in the ambient conditions. A certain volume of sound is necessary to ensure that recipients perceive the sounds of predicted situations in their living environment. Hence, it is required to maintain the volume of natural sounds at one level. If the processing effects used in recording and amplifying the performance alter the sound quality, the recipients might inaccurately recognize a situation such that it differs from the situation they might have perceived in an ordinary environment. Hence, it is necessary to maintain the acoustic quality.

We do not utilize live-broadcasted audio and verbal queues ("sleep", "bath", etc.) because of the privacy of inhabitants and the pressures on caregivers at a group home. Moreover, it is not easy for recipients such as inexperienced caregivers to perceive the sound or situation from a live-broadcasted audio sequence while carrying out other caretaking activities. According to our caregivers who participated in our study, the repetition of verbal queues is a type of alarm sound. They feel the pressure to make some arrangements urgently even in a non-emergency situation.

3. Prototype and User Study

As a first step to examine our system in a group home, we conducted operation and performance tests of each unit in our laboratory (see Fig. 3). We used the software HUGIN (Hugin Expert, Ltd.) to construct the Bayesian network. We used an active RFID (radio frequency identification) system (RF Code Co., Ltd.) to detect the user location information. The active RFID tag has a width and height of 6 cm and 3 cm, respectively. The tag is attached on a user's belt as a key ring. The elderly people have a strong tendency to peel of any objects that are installed on them. We consider that they carry the tag as a lucky charm in a field test of a group home because they are familiar with such charms. Active RFID receivers are installed in locations where activities occur. If an active RFID receiver receives a signal emitted by an active RFID tag, we assume that there is a tag in the vicinity of the receiver and its position is detected. The active RFID receivers are installed at the locations indicated by alphabets ((a)-(e)) in Figure 3. In order to understand an object's status, we use some types of sensors (active RFID and acceleration sensors). For example, acceleration sensors are attached to the doors of refrigerators and store cupboards for detecting their opening and closing motions.

3.1 Performance of Situation Prediction Unit

We investigated how the situation predicted by the situation prediction unit corresponds to the correct situation of a target for a particular period of time. The subjects were three students from our laboratory. During the experimental period of three days, they recorded their location and situation every 30 min. The information of the recorded location and situation must be accurate. The situation prediction unit used the Bayesian network shown in Figure 4. The correlation

Table 1 Correlation between locations and situations in the test

Location	Situation
Refrigerator	Eating or interested in eating
Desk	Some type of work (use of PC, reading, etc.)
Sofa	Coffee break and chat
Private room	Outside a dining room

between the location and situation is shown in Table 1. Such a Bayesian network formed the basis of each subject's daily life. The result of a comparison between the correct situation data sets and the predicted situation is shown in Figure 5. The average percentage of correct predictions for all the subjects is approximately 89%. We found that the situation prediction unit was effective in estimating their situations in this experiment.

3.2 Performance of Notification Unit

3.2.1 Efficacy of Natural Sounds

We investigated whether sound cues could notify a target's situation to an individual. We recruited ten subjects in their twenties from our university. The task for each subject was to listen to sound cues corresponding to seven situations prepared beforehand and select the situation that he/she recognized from the cues. In this test, natural sounds were used as sound cues. In addition, they were required to categorize the sound cues as easy or difficult to recognize. The result is shown in Table 2. The rate of answering correctly varies considerably from one sound cue to another. Although the efficacy of natural sounds is not clearly confirmed, we observed that there is a relationship between the rate of answering correctly and easily recognizable sound cues. Hence, it is important to conduct a survey on the selection of sound cues in advance and use easily recognizable sound cues.

Table 2 Sound cues used in the test and rate of correctly recognizing the notification

Situation	Sound cue	Correct rate
Eating a meal	Use of chopsticks	71.4
Desk work	Typing sound	85.7
Coffee break	Pouring water	57.1
Outside a dining room	Closing a door	100.0
Enter a dining room	Ring sound of doorbell	100.0
Sleep	Snoring	100.0
Driving	Engine sound	100.0

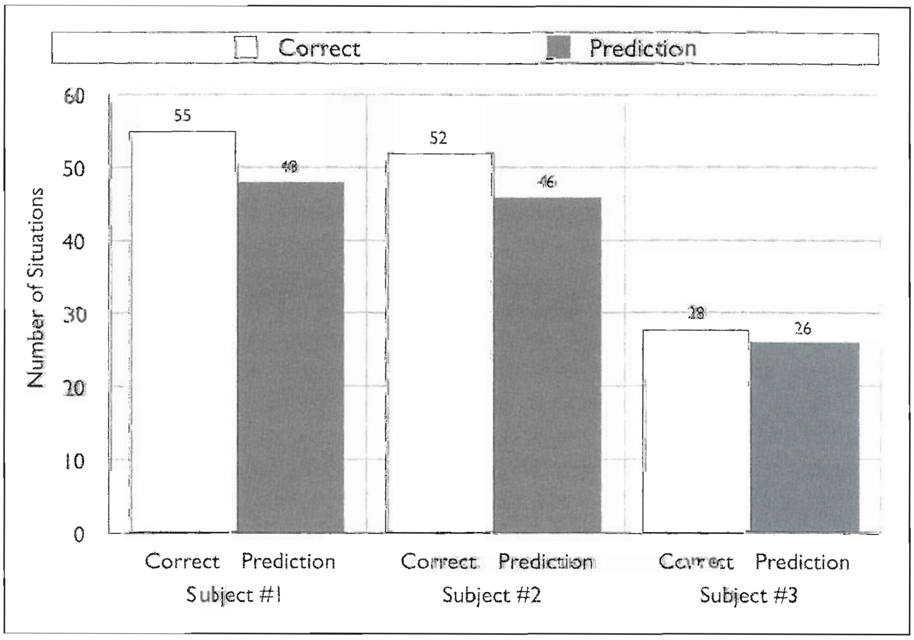

Fig. 5 Result of situation prediction

3.2.2 Efficacy of Natural Sounds with an Acoustic Noise

We investigated the efficacy of natural sounds combined with an acoustic noise. The task of a subject was to listen to a natural sound at a volume based on an arbitrary confidence level of the Bayesian network and identify the confidence level. The other task was to listen to a natural sound with an acoustic noise based on an arbitrary confidence level and identify the confidence level. Both tasks were performed under environments that involved day-to-day noise (TV is switched on) or silence (TV is switched off). Here, the confidence levels employed were 10%, 30%, 50%, 70%, and 90%. The sound of "closing a door" was used as the natural sound because its recognition rate was the highest among the natural sounds used in the previous investigation. A hissing noise was used as the acoustic noise because a certain level of hissing noise can be added to a sound regardless of the sound content. The subjects in this test were the same as those in the previous investigation. In the prototype system, a sound level of 0.0 dB[a] was assigned to the hissing noise for which the confidence level was 0%. The decrease in the level of hissing noise was –0.3 dB for every increase of 1% in the confidence level. Consequently, when the confidence level was 100%, a sound level of –30 dB was assigned to the hissing noise. The result of this investigation is shown in Figure 6. Under a silent environment (TV is switched off), the rate of answer-

[a] Here, 0.0 dB implies the default value when Java Audio Engine is used as the mixer in the Java Sound API.

Table 3 Lower confidence limit for notification

Day-to-day noise	NO		YES	
Acoustic noise	No	Yes	No	Yes
Confidence (%)	45	12.5	47.5	10

ing correctly for the natural sound alone (80%) is evidently higher than that for the natural sound with the hissing noise (10%). We consider that the hissing noise obstructs the subject's hearing under a silent environment because it is particularly noticeable in a silent environment. On the other hand, under an environment that involves day-to-day noise, the rate of answering correctly for the sound cue alone (35%) is marginally greater than that for the sound cue with an acoustic noise (30%). We found that the use of natural sound with an acoustic noise was not effective in this experiment.

As an additional test, we investigated a range of confidence levels when a natural sound is transmitted with an acoustic noise. In this test, a subject listened to a natural sound with an acoustic noise corresponding to an initial confidence level of 100%. The confidence level decreased by 10% each time the subject heard the sound cue. The subject identified the confidence level at the instant he/she experienced a difficulty in hearing the sound cue. This confidence level corresponded to the lower confidence limit of the notification. Further, we conducted the same test for an alteration in the volume of the natural sound. The result is shown in Table 3. In both the environments, the natural sound with the acoustic noise is transmitted at a lower confidence level than the natural sound alone. This suggests that the use of a noise is more effective than the use of a change in the volume of sound for notifying the confidence level.

Based on the above investigations, we consider that the use of an acoustic noise depends on the usage environment. That is, in the case of silent environments such as midnight, the use of only natural sound is more appropriate. In the case of environments that involve day-to-day noises, such as daytime, the use of noise is more suitable for notifying the predicted situations with a confidence level.

4. Conclusion

In this study, we have proposed an audio notification system for the real-time situation of persons in a group home environment. The main contributions of our study are 1) the implementation of a notification system that promotes the awareness of a person's real-time situation by means of sound cues and 2) a testing of the system in laboratory conditions as a first step before evaluating it in a group home. The testing showed that the correct prediction of the subject's situation is approximately 90%. Further, it is shown that the sound cues should be selected according to their environmental dependence. In the case of silent environments such as midnight, the use of only natural sound is more appropriate. In the case of environments that involve day-to-day noises, such as daytime, the use of a noise is more suitable for notifying predicted situations with a confidence level. As a part of a future study, we will test the entire system and improve it for practical use in a group home. In order to select more appropriate sound cues, we will assess the entire environment in which the system is installed.

Acknowledgment
Our study is partly supported by a fund from the Ministry of Education, Culture, Sports, Science and Technology of Japan under the name of Cluster for Promotion of Science and Technology in Regional Areas.

References

1. The Japanese Ministry of Health, Labor and Welfare. White paper on aging society 2006. Available from: http://www8.cao.go.jp/kourei/english/annualreport/index-wh.html
2. Kilander F, Lönnqvist P. A whisper in the woods? An ambient soundscape for peripheral awareness of remote processes. In: ICAD'02: Proceedings of the 2002 International Conference on Auditory Display. 2002. 5 pages.
3. Butz A, Jung R. Seamless user notification in ambient soundscapes. In: IUI'05: Proceedings of the 10th International Conference on Intelligent User Interfaces. ACM; 2005. pp 320-322.
4. Barrington L, Lyons J M, Diegmann D, Abe S. Ambient display using musical effects. In: IUI'06: Proceedings of the 11th International Conference on Intelligent User Interfaces. ACM; 2006. pp 372-374.
5. Russell S, Norving P. Probabilistic Reasoning System, In: Artificial Intelligence: A Modern Approach. 2nd ed. Prentice Hall; 2002. pp 436-470.

Correspondence to:
Hideaki Kanai
Center of Knowledge Science
Japan Advanced Institute of Science and Technology
1-1 Asahidai, Noumi-shi
Ishikawa 923-1292
Japan
E-mail: hideaki@acm.org

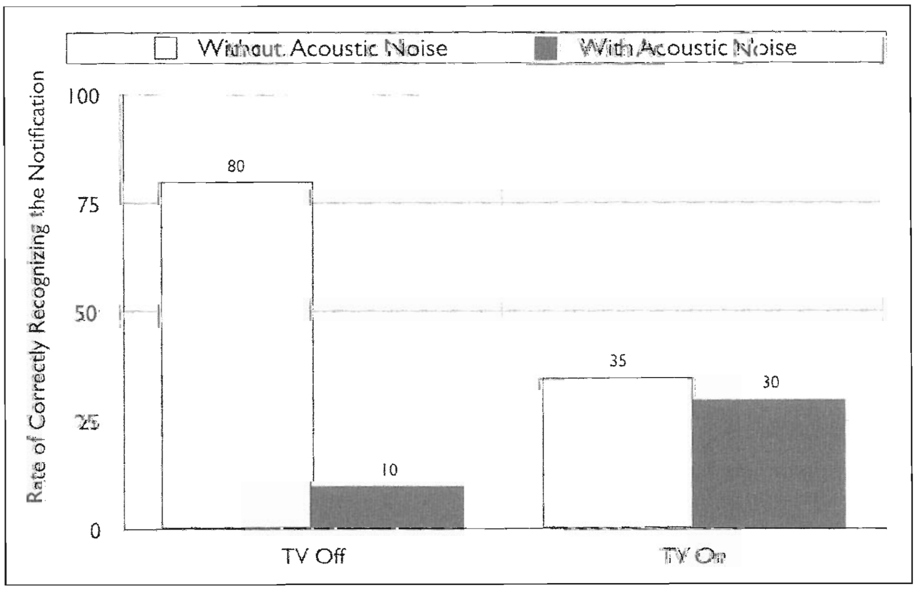

Fig. 6 Result of correctly recognizing the notification