음성인식에 기초한 치매환자 노인을 위한 대화시스템

김성일 · 김병철*

Dialog System based on Speech Recognition for the Elderly with Dementia

Sung-Ill Kim · Byoung-Chul Kim*

요 약

본 연구는 치매노인환자의 생활의 질을 향상시키기 위한 대화시스템의 개발에 목표를 둔다. 제안된 시스템은 주로 세가지 모듈, 즉, 음성인식, 시간테이블에 의해 구분된 대화 데이터베이스의 자동검색, 그리고 간호사의 녹음음성으로 이루어진 맞장구 등의 긍정적인 대답, 등으로 구성되어 있다. 첫 단계로서, 치매환자가 간호시설에서 자주 발화하는 대화의 내용을 조사하였다. 다음으로, 환자들의 요구를 충족시키기 위해 그들의 발화 음성을 자동인식하도록 구성하였다. 여기서 시스템의 응답은 전문 간호사의 녹음음성으로 설계되었다. 시스템의 평가를 위해서 시스템이 도입되었을 때와 되지 않았을 때의 비교연구를 실시하였고, 치료 전문가(occupational therapist)들이 비디오 촬영을 통해서 남성 대상자의 반응을 평가하였다. 평가 결과는 치매환자의 요구를 충족시키는데 있어서 대화 시스템이 전문간호사들보다 더 응답적이었다는 것을 보여준다. 게다가 제안된 시스템은 상호 대화에 있어서 간호사들보다 환자가 더 많이 말하도록 유도함을 알 수 있었다.

ABSTRACT

This study aims at developing dialog system to improve the quality of life of the elderly with a dementia. The proposed system mainly consists of three modules including speech recognition, automatic search of the time-sorted dialog database, and agreeable responses with the recorded voices of caregivers. For the first step, the dialog that dementia patients often utter at a nursing home is first investigated. Next, the system is organized to recognize the utterances in order to meet their requests or demands. The system is then responded with recorded voices of professional caregivers. For evaluation of the system, the comparison study was carried out when the system was introduced or not, respectively. The occupational therapists then evaluated a male subjects reaction to the system by photographing his behaviors. The evaluation results showed that the dialog system was more responsive in catering to the needs of dementia patient than professional caregivers. Moreover, the proposed system led the patient to talk more than caregivers did in mutual communication.

Keywords

Dialog System, Speech Recognition, Dementia, Caregiver, Virtual Conversation,

*중국 청화대학 음성기술센타,

*밀양대학교 컴퓨터정보통신공학부

접수일자 : 2002. 10. 18

I. Introduction

In the recent years, as it has become an aging society both in developed and in developing countries, the overall quality of life(QOL) of the elderly is getting more and more important issue. Particularly, recovering a good health physically and mentally from diseases is an essential purpose of the life for the elderly people. "Therefore, such related technologies and secure environments have been designed for independent living and social participation of older persons suffering from several kinds of diseases. For one of the methods realizing them, the mutual interaction based on conversation has been studied for an increase of the mental health-related QOL and rehabilitation.

This study aims to improve the QOL of the elderly with a dementia[1,2,3]. The dementia possesses unique features that make nursing particularly burdensome. The mental exhaustion in the course of care at home often leads to problems in terms of caregivers mental and physical health, because of the behavioral problems of their family member with dementia. Therefore the QOL of family caregivers is inevitably affected. Namely, family caregivers experience mental stress, burden, and depression as outcomes of nursing activities. Therefore, this study also aims to improve the QOL of family caregivers by lightening their nursing loads to some degree in long-term care[4,5].

In this study, on the basis of these social backgrounds, we have developed the dialog system[6,7,8,9,10] of recognizing and understanding demands of dementia patients mainly depending on the techniques of speech recognition. The system was designed to be a good conversational partner to dementia patients whenever they need. As a consequence, the emotional stability might be recovered by mutual communication, thus having effects of rehabilitation. Furthermore, the system might be helpful for the nursing works of family caregivers at home or professional caregivers at nursing facilities. In order to realize it, our approach is focused on a natural interaction between the dementia patient and the dialog system through spontaneous speech, without using any other interfaces such as a keyboard or mouse, etc. Moreover, the proposed system should be optimized to only one dementia patient because of the diverse symptoms of dementia. In this study, therefore, subject has been confined to one aged person with a severe dementia. However, the proposed system has an extended function that can be applied to other patients with different symptoms of dementia.

II. Outline of Dialog System

In order to realize communication between dementia patient and dialog system, the system for dementia patient is required to be equipped with the following major functions.

(1) The techniques of the command speech recognition were used for speech recognition module. It is because the dementia patients speak the limited and repeated utterances.

(2) For a response of system, it is organized to reply relevantly by speech synthesis or recorded voices to users demands. In this case, a graphical interface with a virtual caregiver is synchronized with the response.

(3) System provides functions of chiming with user or making agreeable responses to demands. This function enables system to interact with user more smoothly and naturally.

(4) System captures the time when the incoming voice signals are given to speech recognition module. At next step, it searches the most suitable response out of database at the

moment, which has been classified into the time schedule of nursing facility.

(5) In case user asks back, system is designed to make the same response as the previous one, because dementia user has a tendency of reconfirming the answers at a frequent rate.

(6) While system is responding, the incoming signals are restrained to prevent it from becoming oversensitive.

(7) System has adaptability to switch day-care service to different environments such as short-stay service of nursing facility or home-care service. It is easily realized by substituting for the desired dialog database suitable for the new situation.

(8) System is adaptable to different users with different symptoms of dementia by registering their individual utterance patterns.

Figure 1 shows the main frame of dialog system with the functions mentioned above.

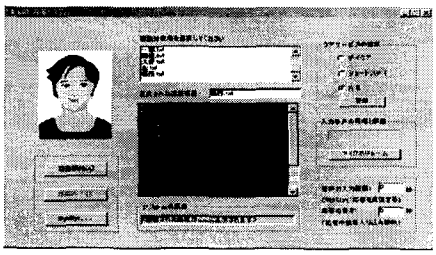


Fig. 1. Main frame of dialog system with several functions such as speech recognition, selective options for both different services and different users, as well as graphical user interface

The proposed system chiefly depends on the function of speech recognition for interaction. However, its high performance is not always guaranteed owing to the ambient noisy environments and slurred speech of the elderly with dementia. In this study, therefore, the additional functions, such as an automatic search of database classified by a timetable of nursing home or chiming with user, supplement a shortcoming of speech recognition for maintaining natural interaction. Therefore, this enables system to promote natural conversation between system and user, even when system fails in recognizing speech of user correctly, or when the incoming speech is not registered in candidate lists. As a result, user might feel as if dialog system hears and understands users words attentively, just like familiar caregivers or family members.

A block diagram is shown in figure 2 that illustrates how conversation can be build between dementia patient and dialog system.

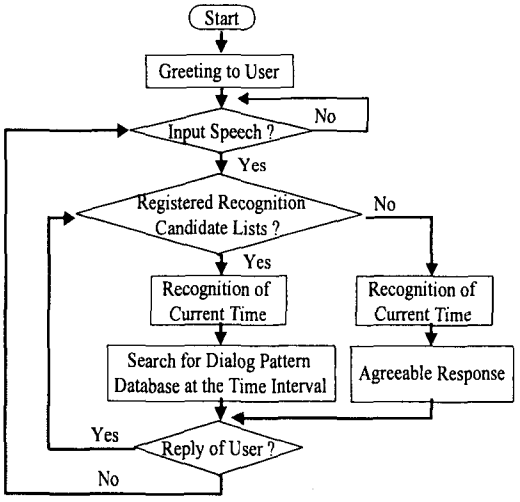


Fig. 2. Flowchart of building virtual conversation between dialog system and dementia user

In case user speaks to system, as a first step, both the incoming speech and the current time are detected by speech recognition engine module and input signals are then compared with all available hypotheses prescribed in candidate list. In the next step, system retrieves dialog database at the corresponding time interval in case speech recognition is done successfully. As a result, it picks out the most likely and suitable response out of the database and then responds with

recorded voices, synchronizing them with a lip of virtual caregiver image. If system fails in recognition, on the other hands, it makes agreeable responses by chiming with user. Moreover, in case input time is acquired within a certain preset time interval, system regards it as a reconfirmation of user to its response. Therefore, system makes the same response as the previous one. The operating system is then ready for detecting a next input speech repeatedly.

III. Evaluation

1. Subject description

The behavioral patterns of a dementia patient were examined at nursing facility for a relevant design of dialog system. For this study, the official approvals from the ethics committee of nursing facility as well as from family members of subject were first obtained in advance. The subject was 72 years old male patient with a vascular dementia. He has received the day-care service at nursing home from 8:30 a.m. to 3:30 p.m. on weekdays.

Since dementia patient like subject is very sensitive to changes in resident environment or life schedule and dementia usually has a disorder of memory, these factors can cause an extreme disappointment or frustration to the patient, occasionally developing an agitation or mental depression. Accordingly, it should be taken into account that the proposed system is designed in order to become a conversational partner as a good listener, sometimes leading him to talk, so that it might help him to restore a comfortable emotional state.

2. Collection of Dialog Database

The frequently used words of subject were surveyed for building dialog database of system. In general, the elderly with dementia often finds it hard to remember the meaning of words used in their daily lives, or to think of words they want to say. In this study, therefore, we selected the most relevant dialogs as to how to talk to dementia patient for a virtual communication between patient and system, according to the following criteria on the basis of the experiences of professional caregivers or occupational therapists.

First, dialog system should try to use words that do not make agitation or frustration worse for dementia patients. Second, system should be designed to call by name or intimate nickname. Third, system is necessary to be designed to speak slowly and distinctly, and use familiar words and short sentences mixed with regional dialects that are familiar to the patients. Fourth, system should be designed to keep things positive, for instance, offering positive choices like "Let's take a rest by the time your family comes here", or "Lunch would be served with your favorite things". Finally, system should be designed to ask simple questions that can be answered with yes or no or one-word answers, when it is difficult to know what they want.

Under these criteria, several sorts of dialog patterns were collected by two female professional caregivers, in which the voices were particularly more responsive for subject than the voices of others. When building the database, more importantly, the changes of patients demands in time should be taken into account to make him retain emotional peace. Moreover, system needs to have a supplementary function of announcing necessary information on the schedule of day-care service.

3. Operating Experiments

In the preliminary investigation, it was noticed that subject demanded something more frequently after lunch. Therefore, the dialog system was set

after lunch when he was in the most uncomfortable emotional state. In experiments, we used recorded voices of caregivers for natural interaction, which had been proved to be more responsive for him than using synthesized ones. The operating experiments of system were performed in the main hall of nursing facility, which was a relatively noisy because it was the place where the nursing home residents have rests or meals. Figure 3 shows a virtual conversation between dementia patient and dialog system using a wireless microphone(WT-1110, TOA, Japan) for free behaviors of subject.

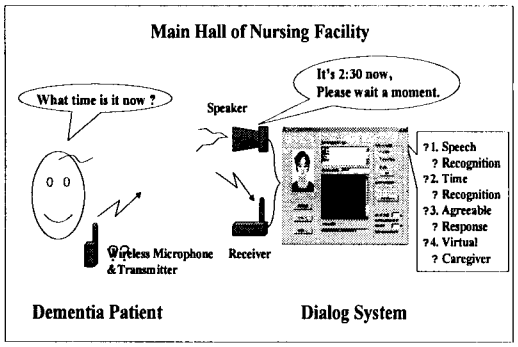


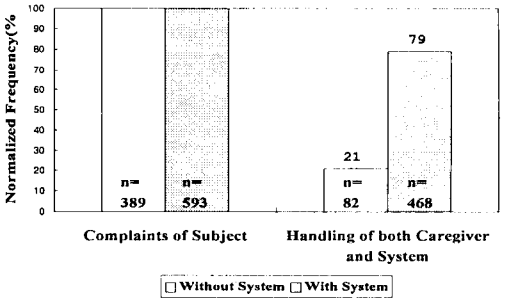
Fig. 3. Virtual conversation between dementia patient and dialog system in the main hall of nursing facility

The comparative study was examined where one of the evaluations was performed with dialog system for 90 minutes of each day during 5 days. In this case, the interruption of caregivers was restricted during the experiment. The other evaluation was performed during the other 5 days without system, where caregivers were allowed to do nursing activities freely whenever subject demanded. Each evaluation test was conducted every second day. During the period of the experiments, occupational therapists observed subjects reactions to system and caregivers, respectively, photographing his behaviors using a video camera.

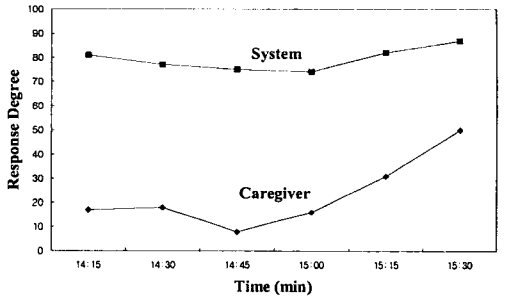
4. Results

For comparative analysis, we examined the frequency of demand of subject, and the frequency of each handling to demand for system or caregivers, as well as the corresponding reactions of subject. The experiments were investigated for 90 minutes of each day during 5 days.

For a quantitative evaluation,"@we induced the handling degree as a ratio of each response frequency for caregivers or dialog system to demand frequency of subject. Figure 4 shows the comparison of handling degrees for both caregivers and dialog system to demand of subject. It was found that caregivers showed the handlings by 21% to demand of subject, whereas dialog system showed 79%.



(a) Change of handling(or response) degree in frequency



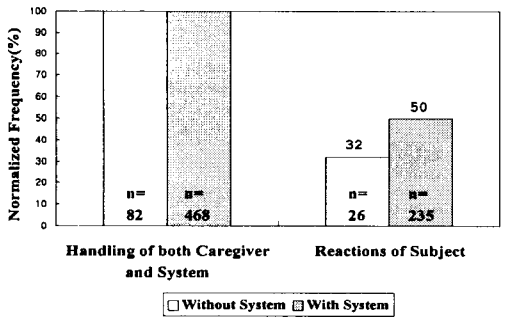
(b) Change of response degree in time

Fig. 4. Comparison of response degrees of both caregivers(namely, the case without dialog system) and dialog system, respectively, to demand of subject

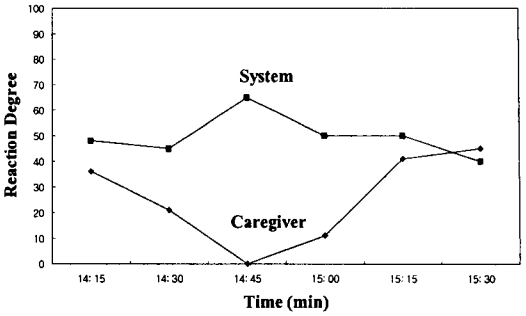
Judging from the comparison results, although caregivers are possible to handle subjects demends by 21%, they were not still responsive by the other 79%. On the other hand, the response of system reached 79% in which it had only the other 21% of non-response owing to the insensitivity of detecting incoming speech signals. In comparison of both graphs, consequently, it revealed that system was more responsive by 58% than caregivers.

Figure 5 shows the comparison of reaction degrees of subject to each response of both caregivers and dialog system, respectively. The reaction degrees were induced by a ratio of each response frequency for caregivers or dialog system to reaction frequency of subject. When system was not given, subject showed reactions by 32% to response of caregivers. When system was given, on the other hand, he showed 50% to response of system. Thus, It was shown that subject was more interactive with system by 18% than with caregivers.

Although the above results show that subject had more frequent interaction with system, it never indicates that more natural conversation is always built. To evaluate how effectively system can maintain a smooth interaction with dementia user, therefore, it is important to study what kind of emotional states (for example, emotional burden or peace) user has in the process of interaction with dialog system.



(a) Change of reaction degree in frequency



(b) Change of reaction degree in time

Fig. 5. Comparison of reaction degrees of subject to response of both caregivers and dialog system, respectively

The professional therapists first examined subjects reactions to each response of both caregivers and system by analyzing the contents of videotapes, as a further analysis of Figure 5. The reactions of subject were then divided into several components. As a first component, subject showed reactions of asking back to get or reconfirm some agreements, particularly when he heard necessary information from caregivers or dialog system as information provider. As a second component, he showed reactions of asking back twice, particularly when the answers to his question were dissatisfactory or different from his expectation. Finally, he showed affirmative reactions with simple answers such as yes or I see. By means of this analysis, we could obtain the frequencies of three different reactions such as affirmation, asking back once or twice, which were indicated in table 2.

Figure 6 shows the comparison of each normalized frequency of three reactions. The notable differences were not shown in the components of asking back once and affirmation. On the other hand, subject had no any reactions in the component of asking back twice to response of caregivers. However, he showed reactions of asking back twice to response of

system. The reactions were shown particularly when the conversation with subject was not built smoothly or naturally, causing uneasiness or an emotional instability.

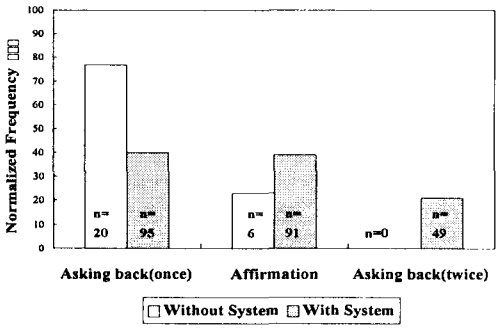


Fig. 6. Comparison of each normalized frequency of three different reactions (n= frequency)

IV. Discussion

In evaluation results, we could find that the accuracies of speech recognition were degraded owing to subjects slurred voices mixed with regional dialects as well as his characteristic accents. Nevertheless, the supplementary functions, such as chiming with user by making agreeable responses, keep conversation smooth, so that he might regard system as a good listener. As shown in comparison results in Figure 4, system was more responsive in dealing with subjects demands or complaints than caregivers of busy nursing schedules. In addition, he had more active interaction with system than with caregivers as shown in Figure 5. Therefore, system might be expected to lighten the loads of nursing works at nursing facilities or at home, by introducing system during their busy time, and thus the QOL of them might be improved as subsidiary effects. Moreover, the rehabilitative effects might be achieved through a prompt response whenever dementia patients need conversation.

In this study, system was optimized to only one dementia patient because of the diverse symptoms of dementia. However, the proposed system has been already equipped with extensional functions to be adaptable to different dementia users by registering their individual utterance patterns. Besides, system has been designed to be flexible to different environments at nursing facility or at home.

However, it is found that we still have essential issues to solve. As illustrated in Figure 6, subject had no any reactions of asking back twice to caregivers because interaction between them was natural and satisfactory. It implies that the natural and timely conversation never gives some kind of emotional stresses or such burdens to dementia patients. When dialog system was introduced, on the other hand, it was noticed that there were occasional reactions of asking back twice. It was mainly due to the failure in recognizing subjects speech correctly, thus causing an unnatural conversation between them. It eventually developed an agitation or uneasiness so that he became stressful or depressive in interaction with system.

Moreover, there is one thing to remember from the experimental results. It is about roles of nonverbal communication such as emotional sympathy, body touch, etc., especially when talking with patients suffering from dementia. Actually, caregivers often talk with their patients while taking hands. Those nonverbal interactions also build a new type of mutual communication as well as human speech.

V. Conclusion

The present study aims to improve the QOL of the elderly who have been suffered from dementia. For realizing this purpose, we have

developed dialog system to respond with a natural dialog to their needs. As a result of the survey, we were able to draw two different conclusions. The one thing was that the proposed system was more responsive in catering to needs of dementia user than professional caregivers. The other thing was that dialog system caused more utterances to dementia user than caregivers did. Moreover, the proposed system has shown a possibility to lighten nursing loads or burdens of caregivers at nursing facilities or at home.

References

[1] Tomoeda CK. Bayles KA. Trosset MW. Azuma T. Mcgeagh A., Cross-Sectional Analysis of Alzheimer Disease Effects on Oral Discourse in a Picture Description Task, Alzheimer Disease & Associated Disorders. 10(4): 204-215, 1996.

[2] Minthon L. Edvinsson L. Gustafson L., Correlation Between Clinical Characteristics and Cerebrospinal Fluid Neuropeptide Y Levels in Dementia of the Alzheimer Type and Frontotemporal Dementia, Alzheimer Disease & Associated Disorders. 10(4): 197-203, 1996.

[3] Fox, Patrick, From Senility to Alzheimer's Disease: The rise of the alzheimer's disease movement, The Millbank Quarterly 67: (1) 58-102, 1989.

[4] Kitwood, T., Experience of Dementia. Aging and Mental Health, 1:13-22, 1997.

[5] K. John, The experience of dementia: a review of the literature and implications for nursing practice, Journal of Clinical Nursing 5: 275-288, 1996.

[6] S. Nakagawa, A. Kai, T. Itoh and S. Kogure, "Speech recognition and understanding of spoken dialogue", Proc. Int. Symposium on Spoken Dialogue, pp.5-1-5-6, 2000.

[7] H.Fujisaki, K.Shirai, S.Doshita, S.Nakagawa, etl., "Overview of an intelligent system for information retrieval based on human-machine dialogue through spoken language", Proc. Int. Conf. Spoken Language Processing, pp.I-70-73, 2000.

[8] Rainer Stiefelhagen, Jie Yang, Alex Waibel, Tracking Focus of Attention for Human-Robot Communication, IEEE-RAS International Conference on Humanoid Robots - Humanoids 2001, 22-24, 2001.

[9] A. Stolcke, K. Ries, N. Coccaro, E. Shriberg, etl., Dialogue Act Modeling for Automatic Tagging and Recognition of Conversational Speech, Computational Linguistics 26(3), 339-373, 2000.

[10] H. Asoh, T. Matsui, J. Fry, F. Asano, and S. Hayamizu, A spoken dialog system for a mobile office robot, Proc. of Eurospeech'99, pp.1139-1142, Budapest, September, 1999.