Dialogue Support for Memory Impaired People

Luca Bellodi*, Radu Jasinschi†, Gerard De Haan*†, Murtaza Bulut†
* Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands
† Philips Research, Eindhoven, The Netherlands
luca.bellodi84@gmail.com, radu.jasinschi@philips.com, g.de.haan@philips.com, murtaza.bulut@philips.com

***Abstract*—People affected by the loss of short term memory and cognitive impairment have serious difficulties in communication. This may lead to social isolation and lack of community access, a fundamental key barrier to independence for people suffering from Alzheimer's Disease, the most common form of memory and cognitive impairment.**

We propose Automated Memory Support for Social Interaction (AMSSI), a system that helps memory impaired people with their social interaction. The system provides active support that may help reducing stress level of patients. AMSSI recognizes visitors, determines the purpose of the visit, monitors the dialogue, determines whether the patient needs support, and provides feedback. AMSSI is tailored to patient needs, it has fast computation, full automation, and can be handled by the patient without supervision. The proposed assistive system can be beneficial for improving the quality of life of patients with mild to moderate cognitive impairments. This paper describes the implementation of the first working prototype of the AMSSI system. Validation user tests are still to be conducted.

I. Introduction

Quality of life and independence are impacted by disabilities. The cost of home healthcare services for elderly people is rapidly increasing as the numbers of disabilities increase within the elderly population. Hence, there is a significant need to find cost effective ways to help the elderly maintain their independence, and at the same time reduce caregiver effort.

A common impairment that affects elderly is memory impairment. Within this group there are Alzheimer's Disease (AD) patients who have cognitive impairments in addition to memory impairments, which impact thought, language and executive function. People with mild to moderate AD may have great difficulty functioning without supervision. For example, AD patients have difficulties to maintain their social life while the disease is progressing [1], because their impairments make it really difficult to interact with others. This has been identified as one of the key barriers to their independence. The loss of memory also causes stress that is usually reflected as undesired behavior of patients [2]. It is important to note that all kind of assistive support needs to be tailored to AD patients. Since they have limited capabilities to process information, a specialized way to interact with them needs to be defined.

It has been investigated how speech interfaces can be used to help these patients. Spoken dialogue systems have been used with the objective of helping patients by reducing the cognitive and memory effort required to use those system [3]. Research results show that there is still a need of age-specific adaptation in order to recognize older users' speech reliably [4].

In this paper we describe a novel solution (Automated Memory Support for Social Interaction (AMSSI)) to help patients with mild memory impairments by giving them support during their interactions with others. Our belief is that such a system can enhance the quality of life for AD patients and can reduce the level of stress during the communication. The support includes helping patients to remember and understand better both the visitor and the topic of the conversation during face to face interactions. AMSSI system assists individuals with AD during their social interactions by means of recognizing the visitor, analyzing the context of the conversation, prompting the conversation, and supporting the patient and visitor if needed. The system is designed for the AD patients at the early stages of the disease, when they are still able to operate such devices. In comparison to existing systems, the novelty of our work is in supporting the patient and visitor in a non-intrusive and automated way by providing multi-modal feedback (to both the patient and the visitor) tailored to the patient's needs and emotions.

The rest of the paper is organized as follows. Section II gives an overview of the related work and shows how our approach differs from the existing technologies. Section III introduces and describes the AMSSI solution and implementation. Section IV discusses the system. Section V presents the conclusions and suggests directions for future work.

II. Related Work

A popular approach is to build an ad-hoc environment equipped with multiple sensors, cameras and panels in order to monitor and support the activities of patients and their interactions in the smart environment. The use of such pervasive computing technology is at the core of Mobile Patient Care-Giving Assistant (mPCA) [5], a device that tracks and captures the attention of the patient with a protocol that uses different senses (smell, vibration, music, audio). In addition it acts as a reminder by using audio feedback, and provides an enhanced level of awareness to the patients by notifying events, through an audio interface. The patient is detected with the multiple sensors installed within the house.

Another branch of research is focusing on enhancing the cognitive capabilities of AD patients. For example, Computer Interactive Reminiscence and Conversation Aid (CIRCA) [6] aims at realizing a device to assist AD patients during supervised reminiscence sessions through a multimedia scrapbook with a specially tailored user interface.

Speech-based support has been tried as well, but there is evidence from literature [7] that audio feedback without priming the conversation (by constant use and training under the supervision of a caregiver) may confuse people with cognitive impairments.

In general, the feedback provided to the patient is mainly audio based, but the cognitive impaired capabilities require feedback tailored to the patients' needs. The user interfaces realized in many systems, even when they take into account memory and cognitive impairments, need to be used under supervision of a health-care professional and thus do not focus on the independence of the patient.

Our work differs from the existing proposals in that we propose a system that can be used without any supervision because it is fully automated. We developed a "transparent user interface" [8], into a system that does not need a whole set up environment and that provides a higher degree of portability and interfacing. AMSSI has been developed to fulfil the necessity for a support during

[0]Luca Bellodi is now at ASML, Veldhoven, The Netherlands. This work was conducted at Philips Research for his Master's Thesis at TU/e.

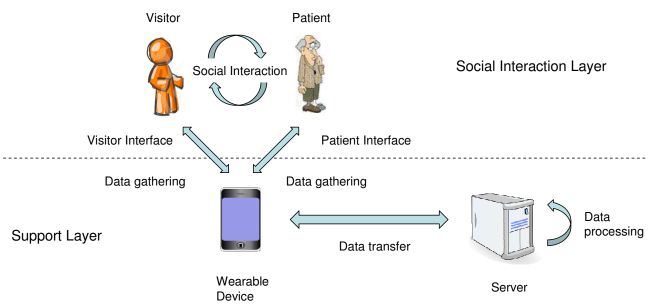

Fig. 1. Automated Memory Support for Social Interaction (AMSSI) system conceptual diagram and information flow. The device gathers data from the interaction, sends it to the server where it is processed and decided if feedback is needed. If that is the case the device provides feedback to visitor and patient.

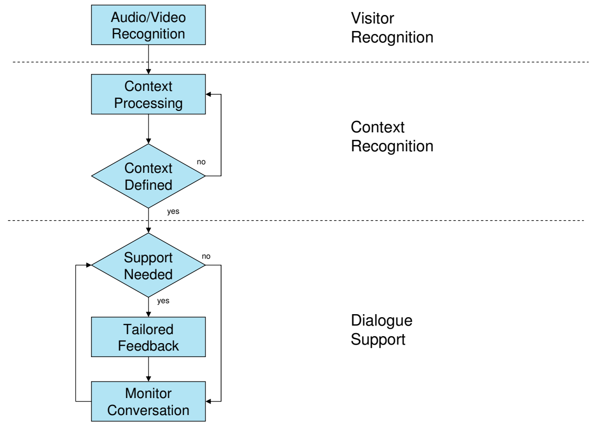

Fig. 2. AMSSI flow diagram: The system recognizes the visitor, identifies the context of the visit, monitors the conversation, decides if support is needed and provides feedback, if necessary.

social interaction, and particularly dialogue management for memory impaired patients. It interacts with a patient in a manner that is suitable to his impairments and, at the same time, aims to reduce the level of stress. Field studies are planned to evaluate its effectiveness.

III. Automated Memory Support for Social Interaction (AMSSI)

AMSSI system can be divided into two parts (see Fig. 1): (a) a social interaction layer, where the patient is interacting with a visitor and (b) a support layer which processes the data obtained from the social interaction layer. The system decides if support is needed and provides it, if necessary. Support can be provided to (i) recognize the visitor, (ii) understand the purpose of the visit and (iii) follow the dialogue.

A common scenario is that the patient meets a friend or receives a visit from a caregiver. It often happens that the patient is not able to recognize the visitor and to remember why he is there. In addition, even if the patient is able to recognize the visitor, he may not be able to follow and actively participate in the conversation due to his memory and cognitive impairments.

Not being able to recognize the visitor or understand the conversation can increase patient's stress level. In this situation, generally, continuing the conversation in a healthy manner is not possible. This can have negative effects on both parties. To address this issue, we have developed a wearable assistive device that is capable of gathering data and providing support to both the visitor and the patient. The device can be used without supervision by the patient himself.

Data collected by the wearable device is wirelessly transferred to the server where it is processed and it is determined if the system needs to provide support. The processed information is sent back to the device that provides the desired support to restore the flow of the patient-visitor dialogue. The support is done by playing a tailored audio or video message helping the patient to understand and continue the conversation. In addition, the device is designed to interact with both patient and visitor, whichever is necessary. A prototype system has been built.

A. System design and implementation

The implemented system comprises different phases (Fig. 2): (i) a preparation phase, where the device sets up the environment to support the expected social interaction; (ii) a visitor recognition phase, where the device detects and recognizes the visitor; (iii) a context recognition phase where the device determines the reason of the visit and decides whether the patient needs initial support or not; (iv) and finally a conversation support phase in order to provide help during dialogue, if needed.

As an example, let us assume that the first scheduled appointment is with Mr. Brown, an old friend of the patient, who is supposed to visit him for drinking a coffee and to have a friendly chat at 2 pm.

B. Visitor recognition

During this mode the device takes pictures every few seconds and runs the face detection algorithm on them. From experimenting we found that a timeframe of five seconds is suitable to take pictures, transfer them to the server and process them. If no face is detected, the picture is stored in a database for further use as a retrospective memory aid to improve recent autobiographical memory in a patient [9]. If a face is identified, then the face recognition starts. At the same time, a speaker recognition module will process the recorded speech of the visitor (not fully implemented yet). Recognition results from both visual and speaker recognition engines are combined to recognize the visitor. The server determines the visitor and the result is sent back to the device.

The device acknowledges the visitor's name and plays the caregiver's voice stating the name of the visitor to the patient (Fig. 2). Coming back to our example, at this stage Mr. Brown is recognized, identified from the device and his name is communicated to the patient. Then the device enters in the context recognition mode.

C. Context recognition

In this phase the objective is to decide whether the patient needs support. The system guides visitor and patient through a sequence of choice-based questions (which require choosing an answer from the provided set). This is a common practice in enhancing communication with AD patients.The information gathered are used to identify the context of the conversation and to decide if the patient is responsive or if he requires support.

As a first implementation of the system we used yes/no questions in order to identify the context of the visit. This method was selected with AD patients' needs in mind, to provide a simple and understandable method of user-device interface. The system asks the patient if he recognizes the visitor, records and processes answers and then it asks if the patient remembers the reason why the visitor is there. In the future versions, speech recognition and natural language processing modules are planned to substitute the yes/no questions and to be used to identify the topic of the conversation. However, considering the patients' memory impairments still remains to be tested what will be the value of this method over yes/no questions.

Based on this information the system decides if the patient needs support and enters in dialogue support mode. For example, if the patient did not recognize Mr. Brown but he remembers that he was waiting for a friend for drinking a coffee, the system decides that the patient is partially responsive. This information is passed to the dialogue support module (Fig. 2).

D. Dialogue support

In this phase the system has to solve the issue of providing support to an unresponsive or partially responsive patient. First step is to provide feedback according to the context recognized in the previous module. Visual feedback has been used, because similar to reminiscence programs, it can support a patient by taking advantage of long term memory [6].

In addition, we wanted to take advantage of the situation where the patient is having a conversation with a responsive visitor, so we can provide feedback to the patient through the visitor as well (Fig. 3). That is why, as soon as the device receives the support command from the server, it alerts the visitor about the situation. This is implemented because the visitor may not be aware of patient's impairments and it may be difficult to understand patient behavior and possible mood changes. As initial feedback for the visitor it has been decided to show colors in order to catch his attention: green for responsive patient, yellow for partially responsive and red for non responsive patient.

Next, the device aims to capture the attention of the patient through vibration [5]. As soon as the patient rotates the device towards himself it senses the orientation change and switches to a patient support mode (Fig. 3). Since patient's attention is captured, the device provides visual feedback, in the form of video or pictures, in order to enhance his cognitive capabilities by stimulating his long term memory. By offering visual cues, the device aims to remind the patient of the current task or topic of conversation or relation with visitor, thus enabling him to better participate in the dialogue [10]. After the context has been recognized and the possible help is provided the conversation can resume.

For example, if the patient is partially responsive according to the previous context recognition phase, the system shows a yellow attention sign and reminds Mr. Brown to pay attention and to be cautious since he is interacting with a partially responsive patient. Then the device vibrates and as soon as the patient turns it toward himself the device shows a video of him and the visitor and their relation in the past (for example watching football together).

At all times the device records the conversation, divides it into chunks of five seconds each of which is sent to the server to be processed and analyzed (Fig. 1) (Five seconds were selected because of their suitability for a fast data transfer and processing). The server recombines the chunks, extracts data and detects whether the mood of the patient is shifting toward a stressed state, in order to prevent or stop it before happening. If a stressed state is recognized, or if the patient is unresponsive the device provides feedback to the patient and the visitor in order to maintain a pleasant conversation.

1) Emotion based assistance: The system provides help based on parameters extracted from the recorded conversation [11]. In order to do this the server gathers data from patient's speech for calculations of energy, unvoiced-to-voiced speech ratio, pitch range and median pitch. From those values it calculates how they are changing during conversation, checking if a heuristic threshold is exceeded. The threshold is adapted (i.e., trained) according to each patient, based on the values collected the first time the patient wears the device. If energy, unvoiced-to-voiced ratio and pitch range are increasing above the threshold values there is a high probability that the patient is approaching a negative mood state that should be prevented. We do emotion classification along valence (positive-negative) and arousal (high-low) dimensions.

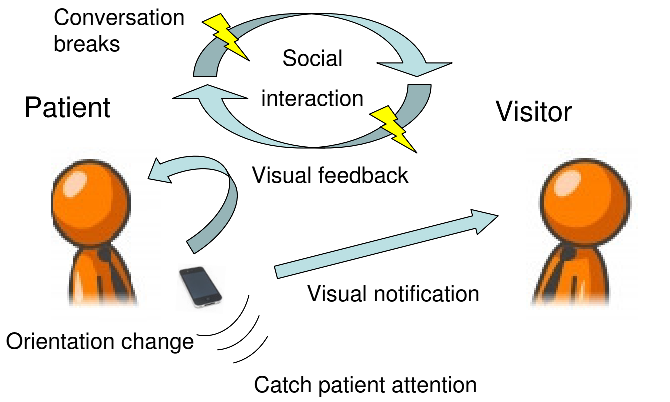

Fig. 3. Different feedback possibilities during the social interaction: (a) visual notification for the visitor and (b) attention catch protocol followed by visual feedback for the patient.

E. Design choices

The data transfer has been implemented according to fast transmission protocols and by sending just small files, by means of dividing the larger ones and recombining them in the server. The server provides the required computational power to process the data and to run the algorithms in a way much faster than it would be possible on the wearable device itself. The visitor recognition phase, for example, helps the patient to recognize the person in front in less than three seconds from the moment the picture is captured.

Next to the fast computation capabilities the system provides an almost automated support, since no input is required from patient or visitor, other than answering a number of yes/no questions at the beginning of the conversation when they first meet. The system is tailored to AD patients by avoiding the use of any demanding cognitive capability, such as having to locate and press buttons.

The system described so far is implemented and a working prototype is built. We used a Nexus One smartphone running with Qualcomm QSD8250 (1GHz clock rate) processor. The assistive application running on the smartphone is developed in the Android OS platform. The server is implemented as a multithreaded C# program which enables connection with the Nexus One phone (worn as a pendant) and the custom software that performs image and speech processing. Wi-Fi and TCP/IP were selected as the wireless connection and communication protocol, respectively.

The face detection is implemented with OpenCV's Face Detector [12]. It is working properly for roughly 95% of photos of people looking straight at the camera. After a face is detected, it is normalized with respect to size, pose and illumination. Then, face recognition implemented with Eigenfaces [13] is run. Preliminary lab test conducted with varying lighting conditions, facial expressions and orientations show a recognition accuracy greater than 90%.

Speech is recorded and processed in continuous segments of five seconds long. Yes/No recognition is implemented simply by detecting the speech (i.e., de-noise and remove silences) and then computing the zero-crossing rate. Using this method, more than 90% accuracy was achieved in lab tests.

The preliminary tests described above were conducted with a limited database of ten people, probably a reasonable assumption for the number of different contacts of an AD patient. Observed face and yes/no recognition errors were mainly caused by changes in environment lighting and noise conditions, respectively. Required processing time capabilities were tested by calculating the average time necessary for the system to interact with the patient. The average waiting time for face recognition was 3.2 seconds after the picture was taken, while the average waiting time for the yes/no recognition module was in the order of milliseconds.

IV. Discussion

The presented study argues that using the described AMSSI system, cognitive impaired patients can be supported during their face to face conversations with healthy people. (It can be extended to support conversations between patients as well.) The system has been built based on the insights and findings from the existing literature explaining needs of patients with cognitive impairments. Based on the literature, we conclude that tools enabling AD patients to recognize people, and to understand, follow and participate in the conversation can increase their life quality.

There is evidence that patients with memory loss commonly lack confidence in company and are generally anxious in everyday life [9]. Our hypothesis is that the subjective levels of anxiety of the patient can be reduced when using the proposed assistive system. The AMSSI can help patients be more confident about their memory, making them able to relax at social events instead of feeling anxious.

Lab experiments conducted with healthy people posing for patients show that a healthy dialogue can be facilitated using the proposed system. Note however that all these claims, the effective impact, and acceptability by users still needs to be verified by field studies with real patients in their own environment. Based on such pilot tests, it may be necessary to speed up the system operation (for example, the face recognition, and speech processing). In addition, it may be necessary to incorporate additional components such as speaker recognition, speech recognition and natural language processing. Use of additional sensors (e.g., accelerometer, light) available in the wearable devices can be also further explored. Since we are supporting the patients who are at the early stages of the disease (i.e., with mild cognitive impairments), we do not think that they will not be able to remember what the device is supposed to be doing. But if this is the case, new methods to help incorporating such devices as a tool in their lives will need to be developed.

V. Conclusions and future work

In this paper a system (AMMSI) which can help patients to communicate face to face with other people is presented. Such a system can be helpful to ease social isolation and difficulties in community access for AD patients [14]. In the long term, the benefits of using the system may be in helping to decrease the deterioration of mental health, maintain a healthy social life and thus self identity [15]. We developed a prototype by combining existing technologies and solutions into a system that is fully automated and tailored to be handled by the patients themselves. The system does not require any expensive set up or smart environment. It is designed in a way that aims at providing helpful support to overcome negative effects of cognitive and memory impairments.

Possible future extension to the system is incorporating wearable devices (e.g. as a bracelet on wrist) that measure physiological sensor data, such as skin conductance, heart rate, heart rate variability and breathing rate. Data gathered by these sensors can then be combined with speech features to more accurately identify emotional state of the patient. Currently the feedback presented to the patient is only visual. We are investigating how it can be effectively combined with audio feedback without disturbing or annoying the patient.

The AMSSI is a working assistive system that has a potential to support memory and cognitive impaired people, such as AD patients, by enabling them to participate actively in face to face conversations. Clinical trials need to be conducted to validate these claims.

References

[1] M. Morris, J. Lundell, and E. Dishman, “Catalyzing social interaction with ubiquitous computing: a needs assessment of elders coping with cognitive decline,” in *CHI ’04 extended abstracts on Human factors in computing systems*, Vienna, Austria, 24-29 April 2004, pp. 1151–1154.

[2] L. Clare, B. Woods, E. Moniz-Cook, M. Orrell, and A. Spector, “Cognitive rehabilitation interventions targeting memory functioning in early-stage Alzheimer’s Disease and vascular dementia,” *Cochrane Database of Systematic Reviews, Issue 4*, 2003, Art. No.: CD003260, DOI: 10.1002/14651858.CD003260.

[3] M. Wolters, K. Georgila, R. Logie, S. MacPherson, J. Moore, and M. Watson, “Reducing working memory load in spoken dialogue systems,” *Interacting with Computers*, vol. 21, no. 4, pp. 276–287, 2009.

[4] R. C. Vipperla, M. Wolters, K. Georgila, and S. Renals, “Speech input from older users in smart environments: Challenges and perspectives,” in *Proceedings of the 5th International on ConferenceUniversal Access in Human-Computer Interaction. Part II: Intelligent and Ubiquitous Interaction Environments*, San Diego, CA, 2009, pp. 117–126.

[5] C. Giraldo, S. Helal, and W. Mann, “mPCA a mobile patient caregiving assistant for Alzheimer patients,” in *First International Workshop on Ubiquitous Computing for Cognitive Aids (UbiCog 02)*, Gothenburg, Sweden, 29 September-1 October 2002.

[6] N. Alm, R. Dye, A. Astell, M. Ellis, G. Gowans, and J. Campbell, “Making software accessible to people with severe memory deficits,” *Proceedings of Accessible Design in the Digital World*, 15-19 August 2005.

[7] M. Fried-Oken, C. Rowland, G. Baker, M. Dixon, C. Mills, D. Schultz, and B. Oken, “The effect of voice output on AAC-supported conversations of persons with Alzheimer’s Disease,” *ACM Trans. Access. Comput.*, vol. 1, pp. 15:1–15:11, February 2009.

[8] A. K. Dey, P. Ljungstrand, and A. Schmidt, “Distributed and disappearing user interfaces in ubiquitous computing,” in *CHI’01 Workshop on Distributed and Disappearing UIs in Ubiquitous Computing*, Karlsruhe, Germany, 1-2 April 2001, pp. 487–488.

[9] S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, and K. Wood, “Sensecam: A retrospective memory aid,” in *Proc. 8th International Conference on Ubicomp*, Berlin, Germany, 17-21 September 2006, pp. 177–193.

[10] B. Wilson and N. Moffat, *Clinical Management of Memory Problems*. Chapman & Hall, 1992.

[11] M. Bulut, S. Lee, and S. S. Narayanan, “Recognition for synthesis: Automatic parameter selection for resynthesis of emotional speech from neutral speech,” in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Las Vegas, Nevada, Apr. 2008, pp. 4629–4632.

[12] G. Bradski, “The OpenCV Library,” *Dr. Dobb’s Journal of Software Tools*, 2000.

[13] M. A. Turk and A. P. Pentland, “Face recognition using eigenfaces,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE Comput. Press, pp. 586–591.

[14] N. H. Narayanan and S. Fickas, “Ameliorating cognitive impairments,” in *CHI’01 Workshop on Distributed and Disappearing UIs in Ubiquitous Computing*, Karlsruhe, Germany, 1-2 April 2001, pp. 11–18.

[15] M. M. Ball, M. M. Perkins, F. J. Whittington, C. Hollingsworth, S. V. King, and B. Combs, “Independence in assisted living,” *Journal of Aging Studies*, vol. 18, pp. 467–483, 2004.