Integrating 3D Facial Model with Person-Centered Care Support System for People with Dementia

Shota Nakatani(✉), Sachio Saiki, and Masahide Nakamura

Graduate School, Kobe University, 1-1, Rokkodai-cho, Nada, Kobe 657-8501, Japan
shota-n@ws.cs.kobe-u.ac.jp, sachio@carp.kobe-u.ac.jp,
masa-n@cs.kobe-u.ac.jp

Abstract. Our research group has been studying a speech communication system with a virtual agent (VA), to support person-centered care (PCC) of people with dementia (PWD). The current system uses an unfamiliar avatar for the VA, which causes a limitation in the care effects. In this paper, we develop a novel system that dynamically creates a VA based on a given facial image. The proposed system constructs a three-dimensional model based on facial landmarks within the image. It then stretches and transforms some portions of the 3D model to generate facial expressions. From just a given picture, the proposed system easily generates a communication agent familiar with individual PWD. Hence, it can implement (virtual, but effective) conversations with familiar partners.

Keywords: Person-centered care · Person with dementia · Virtual agent

1 Introduction

Japan is facing hyper-aging society. The number of people with dementia (PWD) will reach 700 million in 2025, where one-fifth of five elderly people in Japan will suffer from dementia [1]. The person-centered care (PCC) is an ideal care for PWD, which monitors and understand individual circumstances, and plans and executes optimized care. In practice, however, the PCC completely relies on human effort, which poses heavy physical and mental burden on family and caregivers.

To cope with the problem, our research group has been studying a PCC support system for PWD, exploiting the latest IoT and cloud technologies [2]. Gathering and analyzing sensor data from home of PWD, the system understands activities and contexts of PWD at home [3]. The system then generates dialogues, and talks to PWD through the virtual agent technology [4]. The virtual agent (VA) is a human-looking animated chat-bot program operated on a lap-top PC. Using the speech recognition and synthesis technologies, a PWD can communicate with the VA as if (s)he talks to a human partner.

However, the current system uses an artificial avatar for the VA. Hence, the care and advices from unfamiliar avatar does not motivate the PWD very well, which poses limitations on concentration and engagement of the PCC.

This paper develops a new technology that introduces more familiar agent as the VA, in order to achieve more concentrated and effective care within the PCC system.

© Springer International Publishing AG 2018
W. Karwowski and T. Ahram (eds.), *Intelligent Human Systems Integration*,
Advances in Intelligent Systems and Computing 722,
https://doi.org/10.1007/978-3-319-73888-8_34

We generate a three-dimensional model from a given picture of a face of a familiar person of a PWD, and integrate the 3D facial model as the VA of the PCC support system. More specifically, for a given picture of a face, the system extracts landmark points of the face using face recognition algorithm. Based on the landmark points, the system then stretches and transforms some portions of the 3D model, to dynamically generate facial expressions. Also, for a given voice data, the system synchronizes the lip motions, as if the 3D model speaks to the PWD. We integrate the generated model to the PCC system, and uses it as more effective avatar of the VA familiar.

From just a given picture, the proposed method can easily generate a communication agent familiar with individual PWD. Hence, it can implement (virtual, but effective) conversations with familiar partners of PWD, such as close relatives, child, close friends and so on. As a result, the proposed system contributes to the implementation of more effective PCC support system.

2 Preliminaries

2.1 Person-Centered Care Support System with Virtual Agent

Our research group has been studying how the ICT can support a person with dementia (PWD) at home. The concept of person-centered care (PCC) defines an ideal care for PWD, which monitors and understand individual circumstances, and plans and executes optimized care. Our current aim is to provide person-centered communication for PWD using the virtual agent (VA) technology. The VA is a human-looking animated chat-bot program operated on a PC. Using the speech recognition and synthesis technologies, PWD can communicate with the VA as if (s)he talks to a human partner.

In our previous research, we have developed a system, called Virtual Care Giver (VCG), using the VA [1]. VCG was designed to be able to cooperate with Web services, to integrate IoT, smart home and cloud. Based on preferences of a PWD, VCG provide a personalized care for the PWD via VA. Delegating the communication care with VCG, a human caregiver can concentrate on human-centric tasks that cannot be done by ICT.

Our preliminary experiment shows that VCG can achieve useful context-aware care for PWD, including, daily greeting, schedule reminder, and prompting medication.

However, there is also a limitation caused by the looking of VA. Currently, the VA of VCG is implemented by MMDAgent [5], which displays 3D animation character. So, the care and advices from unfamiliar avatar do not always motivate the PWD very well. This poses limitations on concentration and engagement of the PCC.

2.2 Face Recognition in Cognitive Computing

Cognitive computing is a computing paradigm where a system imitates the human brain and learns itself to derive answers. It often refers to the emerging technologies that can analyze non-numerical data, such as language, picture and voice, which were difficult to understand by the conventional computing. Face recognition and analysis are major technologies within the cognitive computing. Recently, several companies publish Web

API for face recognition. For instance, Microsoft provides Face API and Emotion API [6] within the Azure Cloud Services.

Most face recognition technologies (e.g., [7–9]) use machine learning to detect characteristic points on a face such as eyes, nose, mouth and so on. Such points of face are called facial landmarks. When a new picture is given to the machine, it tries to extract the facial landmarks within the picture. Figure 1 shows facial landmarks extracted from my picture using Face++ [10]. Face++ is developed based on [8]. It can derive 83 landmarks as coordinates on a picture. Using these landmarks, we can perform operations to move any specific parts of the face in the picture.

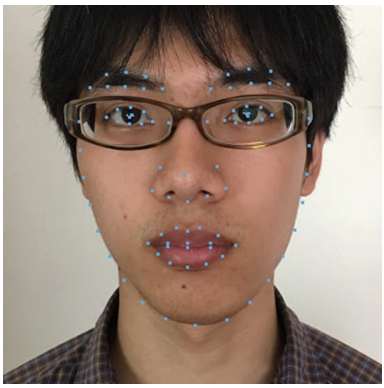

Fig. 1. Facial landmarks derived by Face++

3 The Proposed System

3.1 Key Idea

As a key idea to achieve more interesting communication care for a PWD, we aim to generate a VA with a face of a familiar person of the PWD. Using the face recognition and analysis technologies, we generate a 3D facial model from a picture of the familiar person. We then integrate the generated model to the PCC system, and create a situation where the PWD talks with the familiar person. As a result, we can implement more quality PCC, compared to the conventional artificial VA.

In the following sub-sections, we describe functional requirements of the proposed system, and design of the service API. We also illustrate how the proposed system is integrated with the PCC support system.

3.2 Functional Requirements

Based on the analysis of the previous system, we identified that the proposed system must implement following three key features.

- *Feature F1 (generate a VA):* For a given picture of a face, the system can dynamically generate a VA with the face.
- *Feature F2 (instruct the VA to speak):* For a given text, the system can instruct the generated VA to speak the text.
- *Feature F3 (instruct the VA to change expression):* For a given command, the system can change the facial expression of the generated VA.
- *F1* implements a fundamental requirement of the system to create a VA of a familiar person of PWD. With this feature, every PWD can generate his/her own VA. *F2* implements an ability of speaking in the VA. With this feature, the VA is able to be a conversation partner. *F3* implements an ability of emotion in the VA. With this feature, the VA can behave like a human being, and produce friendly conversation with non-verbal communication.

3.3 Design of Service API

As considered in the previous system, we intend to integrate the VA with IoT and smart home. Hence, we deploy the features mentioned in Sect. 3.2 as Web service. By doing this, external applications within the PCC system can easily use the features *F1*, *F2*, *F3* via Web API. We define three Web APIs that define *F1, F2* and *F3*, respectively.

CreateVA(FaceImageFile). Generates a VA from an image file whose URL is specified by *faceImageFile*. For a given picture of a face, the system extracts landmark points of the face using a face recognition algorithm, and generates a 3D facial model. By default, we suppose that the VA do natural motions such as blink, swaying its face, and so on.

Speak(Text). Instructs the generated VA to speak (read) a text sentence specified by *text*. We suppose, within the API, that voice data (wav, mp3, etc.) should be synthesized from the text using a text-to-speech technology. As the VA plays back the voice data, we also suppose that the VA should synchronizes the lip motions, by stretching lips within the 3D facial model.

ChangeExpression(Expression). Changes the facial expression of the generated VA by a label specified in *expression*. Based on the extracted facial landmark, the system stretches and transforms some portions of the 3D model, which change facial expressions. Typical expressions include normal, smile, angry, sad, surprised, and fear.

3.4 Integrating into PCC Support System

We suppose to deploy the above Web API on a Web server called *VA Manager*. Figure 2 shows a sequence diagram showing how VA Manager works with the PCC support system. Here we assume a scenario that a user creates a VA and greets to it. The user first registers an image file (*"faceImage.png")* to the PCC support system. The PCC support system then executes *createVA*(*faceImage.png)* to generate the VA. The VA looking like the person in *faceImage.png* is displayed on user's PC.

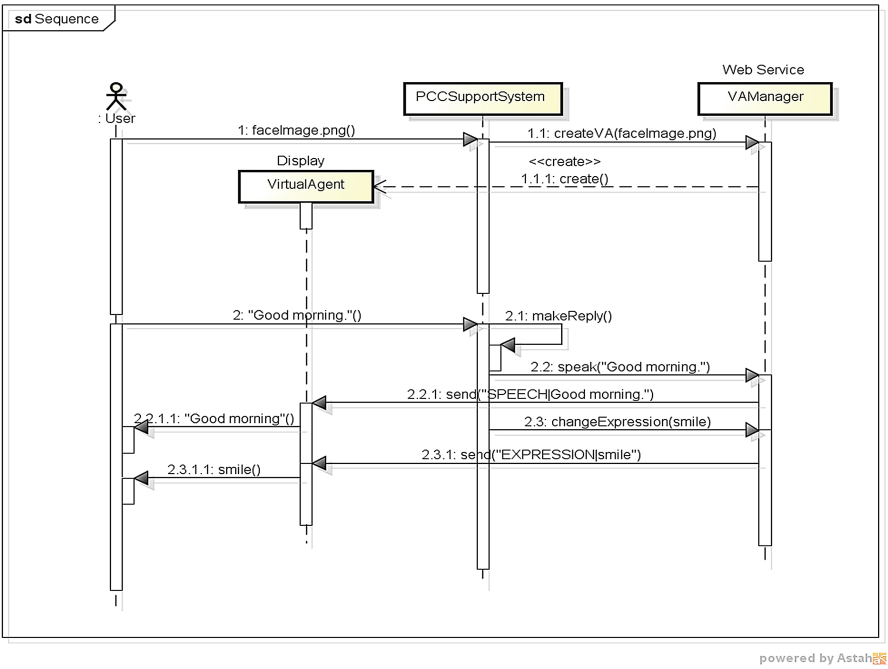

Fig. 2. The sequence of integrated system.

Suppose that the user says to the VA "Good morning". PCC system recognizes the user's voice by speech recognition, generates the reply "Good morning.", and executes speak("Good morning.") to instruct the VA to speak. At the same time, the system executes changeExpression(smile) to instruct the VA to smile. As a result, the VA on the PC screen replies to the user, "Good morning" with smile.

4 Implementation of Prototype

Based on the proposed system, we have implemented a prototype system. Technologies used in the implementation are as follows.

VAManager. Development Language: Java, Web Server: Apache Tomcat, Web Service Frameworks: Apache Axis2.

VirtualAgent (MPAgent). Development Language: C#, MotionPortrait SDK [11].

VAManager was implemented by Java, and deployed on a Tomcat Web server. Based on requests from the PCC support system, it sends commands to VirtualAgent. We implemented the virtual agent using MotionPortrait SDK. It provides powerful libraries for generating and operating 3D facial model. With the SDK, we were able to implement features *F1* to *F3* very efficiently in C#. In the program, we also put a socket to receive commands from VAManager.

Figure 3 shows example facial images in the prototype system. The images show, from left to right, the original face picture, the VA with happy expression, the VA with sad expression, the VA that is speaking.

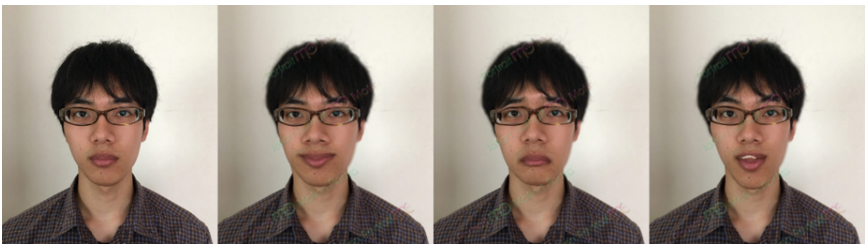

Fig. 3. Images of a virtual agent produced by the prototype system.

5 Conclusion

In this paper, we propose a system that aims more efficient person-centered dementia care with the virtual agent (VA) technology. Using the face recognition technology, the proposed system generates a 3D facial model from a given picture, integrates the model as a personalized VA of the PCC support system. Since every user can easily generate his/her own VA with a preferred face, the system can be expected to achieve more effective PCC. In our future work, we conduct an experiment with actual subjects, to investigate if the looking of VA really affects the performance of the care.

References

1. Cabinet Office, Government of Japan: Annual report on the aging society (2017). http://www8.cao.go.jp/kourei/whitepaper/w-2017/html/zenbun/s1_1_1.html. Accessed June 2017
2. Tokunaga, S., Tamamizu, K., Saiki, S., Nakamura, M., Yasuda, K.: VirtualCareGiver: personalized smart elderly care. Int. J. Softw. Innov. (IJSI) **5**(1), 30–43 (2016)
3. Tamamizu, K., Sakakibara, S., Saiki, S., Nakamura, M., Yasuda, K.: Capturing activities of daily living for elderly at home based on environment change and speech dialog. In: International Conference on Digital Human Modeling 2017 (DHM 2017). LNCS, vol. 10287, pp. 183–194, July 2017
4. Sakakibara, S., Saiki, S., Nakamura, M., Yasuda, K.: Generating personalized dialogue towards daily counseling system for home dementia care. In: International Conference on Digital Human Modeling 2017 (DHM 2017). LNCS, vol. 10287, pp. 161–172, July 2017
5. MMDAgent. http://www.mmdagent.jp
6. Emotion API. https://azure.microsoft.com/en-us/services/cognitive-services/emotion/
7. Taigman, Y., Yang, M., Ranzato, M.A., Wolf, L.: DeepFace: closing the gap to human-level performance in face verification. In: CVPR 2014 (2014)
8. Fan, H., Cao, Z., Jiang, Y., Yin, Q., Doudou, C.: Learning deep face representation. arXiv: 1403.2802

9. Zhang, Z., Luo, P., Loy, C.C., Tang, X.: Facial landmark detection by deep multi-task learning. In: European Conference on Computer Vision (2014)
10. Face++. https://www.faceplusplus.com
11. MotionPortrait. https://www.motionportrait.com