Demo: Talk2Me – A Framework for D2D Augmented Reality Social Network

Jiayu Shu*, Sokol Kosta†, Rui Zheng*, and Pan Hui*‡

*HKUST-DT System and Media Laboratory, Hong Kong University of Science and Technology, Hong Kong
†Department of Electronic Systems, Aalborg University Copenhagen, Denmark
‡Department of Computer Science, University of Helsinki, Finland
jshuaa@ust.hk sokol@cmi.aau.dk rzhengac@ust.hk panhui@ust.hk

Abstract—In this demo, we present Talk2Me, an augmented reality social network framework that enables users to disseminate information in a distributed way and view others’ information instantly. Talk2Me advertises users’ messages, together with their face–signatures, to every nearby device in a Device–to–Device fashion. When a user looks at nearby persons through her camera–enabled wearable devices (e.g., Google Glass), the framework automatically extracts the face–signature of the person of interest, compares it with the previously captured signatures, and presents the information shared by this person to the user. We design a lightweight and yet accurate face recognition algorithm, together with an efficient distributed dissemination protocol. We integrate their implementations in an Android prototype.

I. INTRODUCTION

The continuous proliferation of mobile and wearable smart devices, together with their increasing computational power and multitude of sensors, have contributed to a rapid revolution on how people nowadays perceive the world and maintain social relationships. For example, people use social networks to share their ideas, to stay in touch with other users, etc.

In many situations though, people may be more interested in sharing quick information to people that are physically nearby. Imagine a scenario where a technology workshop is held. Start-ups, investors, experts, and developers are participating. Alice, who runs a startup company, is presenting their products to people around. However, Alice would like to convey a specific information to a certain audience in an efficient way. Meanwhile, Bob is a student looking for a job. He is overwhelmed by the surrounding conversations and presentations.

In this case, an instant and efficient social network for people in the physical proximity is needed. Though various types of online social networks have been used, they only link people that already know each-other, which cannot meet the demand of exchanging information to enhance social interactions with new people in the vicinity. Therefore, we need a new kind of pervasive social network that meets the following requirements and functionalities. First, the network can be created temporarily, without the need of any infrastructure or history records. Second, people can easily modify or update the information they want to deliver. Third, people are able to view and match the information with the person who sends it effortlessly, such as in an augmented reality way.

In this demo paper, we present Talk2Me, a novel framework for Device-to-Device Augmented Reality Social Network (D2D–ARSN) on mobile devices, which allows people to share messages with nearby users in a distributed way, and view others’ shared information instantly [1]. Previous works have already highlighted the usefulness of such distributed social networks. For example, E-SmallTalker connects proximate users based on their shared personal information [2]. Advancing on previous ideas, Talk2Me provides timely and dynamic information sharing and viewing functionalities that facilitate social interactions, especially for initiating conversations and making friends with people in physical proximity. Everyone using the framework is able to shout her thoughts, offering small talk topics or useful messages to people in physical proximity for starting a conversation. Then, all people located nearby that have installed the framework will be able to listen to and view these messages through their camera–enabled devices, such as smart glasses or smartphones. For example, in the scenario shown in Figure 1, Alice broadcasts a message that she is looking for skilled mobile app developers. Bob looks at people around through his camera–enabled device with Talk2Me installed, and sees Alice’s message, so he goes to Alice for further information.

The main contribution of this work is the proposal, design, and implementation of the first, to the best of our knowledge, distributed social network that works in augmented reality, by spreading information with embedded face–signatures.

II. SYSTEM OVERVIEW

In this section, we first introduce the major components of Talk2Me framework. We then describe how Talk2Me works and its implementation.
A. Architecture

We define two roles in an Augmented Reality Social Network (ARSN): promulgator and recipient. A promulgator is the person who broadcasts messages to nearby people. A recipient is the person who receives others’ broadcasted information. Users can be both promulgator and recipient at the same time. Figure 2 shows the architecture of the system. It consists of five modules.

**Input Module:** This module gets the information input from the user. If the message is a text for example, it opens a text field where the user can write the message.

**Camera Module:** This module handles the interface used for i) getting visual data—i.e. capturing frames, and ii) for viewing the information associated with a face–signature. When a promulgator creates his face–signature, or a recipient wants to access the information shared by another person, the camera module is activated. This module captures the photo frames of the physical world and delivers them to the face module for further image processing. Upon receiving results from the face module, it displays them on the screen of a smartphone or smart glasses.

**Face Module:** This module is responsible for face detection, face–signature extraction, and face recognition. It works when a promulgator provides his/her face–signature for message preparation. Moreover, this module is also involved when a recipient captures a frame. It runs a face matching algorithm and retrieves the information associated with the matched signature after face detection and feature extraction. Finally, the information is forwarded to the Camera Module.

**Message Module:** This module processes, prepares, and stores messages. On receiving a message from other devices through the networking module, it first processes the message in order to perform corresponding actions, such as storing the message internally or transmitting new messages. When necessary, this module gets the message from the input module and associates it with the face–signature of the promulgator, and then forwards the messages to the networking module.

**Networking Module:** This module receives messages from the message module of a promulgator’s device and transmits them to the others. In a recipient’s device, it listens for incoming packets, which upon reception are passed to the message module.

B. Design and Implementation

We implement the prototype on Android smartphones. A simple user interface is shown in Figure 3(a).

As a promulgator, she needs to first provide her information that she wants to be advertised. For example in our prototype, a promulgator can input the text such as “CS student”. Later, if she wants to update her information, she just inputs the new text and presses the **Update My Info** button. Next, the promulgator is supposed to create her face–signature, which will be send together with her information. Talk2Me will generate an unique face–signature for the promulgator. The promulgator simply clicks the **Update My Signature** button and looks at the camera of the device. Then Talk2Me will capture several pictures of the promulgator that contain her face. The promulgator can also turn her head slightly and make different facial expressions, so that the face–signature generated can be more robust. After all the preparations are done, the promulgator can press the **Start Service** button to broadcast her information. In Figure 3(a), the Talk2Me is already running. The **Start Service** button has changed to **Stop Service**, which will terminate the information dissemination process once clicked.

As a recipient, he just needs to start the service without proving any information, if he only wants to receive others’ messages. He then presses the **View Information** button to see others’ information. If the person captured by the camera is recognized, the information from this person will be displayed. For example in Figure 3(c), Talk2Me recognizes a promulgator correctly, and displays her information on the screen.

To realize these functions, the most essential parts of the system are the **Image Processing** and the **Dissemination Protocol**. We describe the design and implementation of these two parts briefly.

1) **Image Processing**: It deals with face detection, face–signature extraction, face recognition.

**Face Detection:** We use open source Adaboost cascade classifiers [3] through the OpenCV library [4] to detect face regions in the frame. In addition, we use skin color filters and the Dlib library [5] to reduce the number of false positives, further increasing face–detection precision.

**Face Representation:** A user’s face–signature refers to the encoding of his face. We adopt the state–of–the–art Convolutional Neural Network (CNN) based feature representation, which is a 256–dimensional feature array from a lightened–CNN model [6]. We use the 1KB output of its last fully connected layer as face-signatures. To perform the feature extraction process, we use the Caffe deep learning framework [7] and the Caffe Android library [8].
Face Recognition: We first calculate a similarity score between the extracted signature of the current image frame and the previously received face–signatures from all the users to find the number of matches per person \(C_i, i = 1, 2, \ldots, P\). Based on the number of matches we select two best candidates, whose number of matches are \(C_r\) and \(C_s\). Only if the difference of the matches \(|C_r - C_s|\) exceeds a certain margin, or if a single candidate has matches, we can conclude that the person is recognized. We set the margin as half the number of features per person after cross validation.

2) Information Dissemination: It deals with message transmission and reception. We define three types of messages: 1) HELLO, used for advertising the presence of a user; 2) REQ, used to ask users to transmit their information; and 3) INFO, used to send the information.

The HELLO message consists of the promulgator’s unique ID and his current information version. Upon receiving a broadcasted HELLO message, a recipient’s device will check if it already has this promulgator’s information, by looking at the ID and information version embedded in the message. If no information from this promulgator is stored, the recipient sends a REQ message, which contains the promulgator’s ID and the information version. The promulgator responds to a REQ message by sending an INFO message, which contains ID, information version, the actual information, and user’s face–signature. In our implementation, the information field is limited to 140 bytes, and the features field is limited to 10KB.

III. DEMO SETUP AND REQUIREMENTS

The demonstration setup includes some Android smartphones that have installed Talk2Me. Devices are able to communicate with each other via UDP over Wi-Fi through an access point.

In our demo, people are presented with the application on Android smartphones. They can input the text information and provide their face–signatures using the camera of the smartphone. Then they can start the service to broadcast their messages and receive others’ messages. Later, if the camera of the device captures other users whose INFO message has been received and stored in the device, his/her information will be displayed on the screen associated with the face.

IV. CONCLUSION

In this work, we designed and implemented Talk2Me, the first, to the best of our knowledge, distributed social network framework based on real–time augmented reality. Using Talk2Me, people can disseminate and receive information from people physically nearby in D2D fashion, and then view the information in an augmented reality way. Talk2Me provides a seamless information–sharing channel that we believe will encourage social interactions.

ACKNOWLEDGMENT

This research has been supported, in part, by projects 26211515 and 16214817 from the Research Grants Council of Hong Kong.

REFERENCES