Sign Language Augmented Reality Real-time Animation

Problem Statement

Deaf and mute people (DM) are facing difficulties in their daily activities in different life aspects. They suffer from community's rejection because of their inability to communicate and feel embarrassed or frustrated while trying to communicate with normal people.
Recent studies have shown that more than 360 million patients worldwide are DM (Over 5% of the world's population). 91.1% are over 65 years, while 8.9% are children under the age of 18 have no harmony with normal children of the same age.

Although the growing interest in this area has resulted in many solutions in various fields especially mobile applications motivated by the wide use of smartphones in our daily activities, most of the current available mobile applications seek to serve the deaf people in a very limited way.

For instance, there are some applications that help in communication but has drawbacks due to the discomfort in use.
Sign Language Augmented Reality Animated (SLARA) Application for DM people is proposed, the idea is to help them to communicate normally and easily using SL. An avatar-based translation system from speech to SL for deaf people is introduced. The technique is composed of a database of the captured 3D motions of SL. The SL motion will be captured using camera. In addition, speech is collected by mobile's microphone, the recognized words are translated to signs displayed on glasses.

Methods and Materials

Materials:

SLARA mobile application depends mainly on Unity game engine as it supports most of the plugins, support different platforms as IOS and Android and it is easy to design augmented reality applications with it.

The needed plugins are Google cloud speech recognition plugin (GCSR) that uses Google Cloud Speech to Text API, and Google cloud text to speech plugin (GCTS) that uses Google Cloud Text to Speech API. These plugins are used to translate the detected voice by the normal person into text and the sign language into voice. These plugins need internet connection to connect to Google Cloud APIs. These plugins support many languages up to 14 languages.

We also used the SQLite database in order to map the videos with the text to retrieve them in almost real time. We used the Vufine Wearable Display which allows the user to deal in a hands-free environment by only an HDMI cable which connects the mobile phone to the Vufine Wearable Display.

A small camera is used to track and detect the motion of user's hands and fingers to be processed. This camera is placed on the Vufine Wearable Display as shown in Fig. 1 with an angle of 45 degrees. A built-in microphone with the camera is used for collecting sound to be converted into text using GCSR.

Fig. 1. The Vufine Wearable Display and the camera Prototype

We made an avatar, we called it SLARA, which is a real-life model performs SL videos for the user.

We followed a specific technique to prepare the videos that form our avatar SLARA:

● We captured 2 to 3 seconds videos which presents either a word or expression of the real-life model, then the videos are trimmed using Filmora Wondershare video editing software. A green screen is placed as our character's background because it won't affect lighting of the character if the background is removed.
● Then some visual effects are added to cartoonize the real-life model using Cartoon & Plastic effects in Adobe after effects.
● Unity video player component doesn't accept MOV format, So FFMPEG scriptable converter is used to convert MOV videos into WEBM format, keeping the alpha channel enabled to get the videos with a transparent background.

The mobile application is just an interface for the user to adjust the settings of the application, see the videos of SLARA and close the application as in Fig. 2. We tried to make it easy in use for the perfect experience.

The first image to the left is a loading screen to open the application, the middle image is the screen is the main screen where the avatar is shown in the mobile in addition to the Vufine Wearable Display. The third image to the right shows the side menu where the user can adjust settings and display the SL videos.

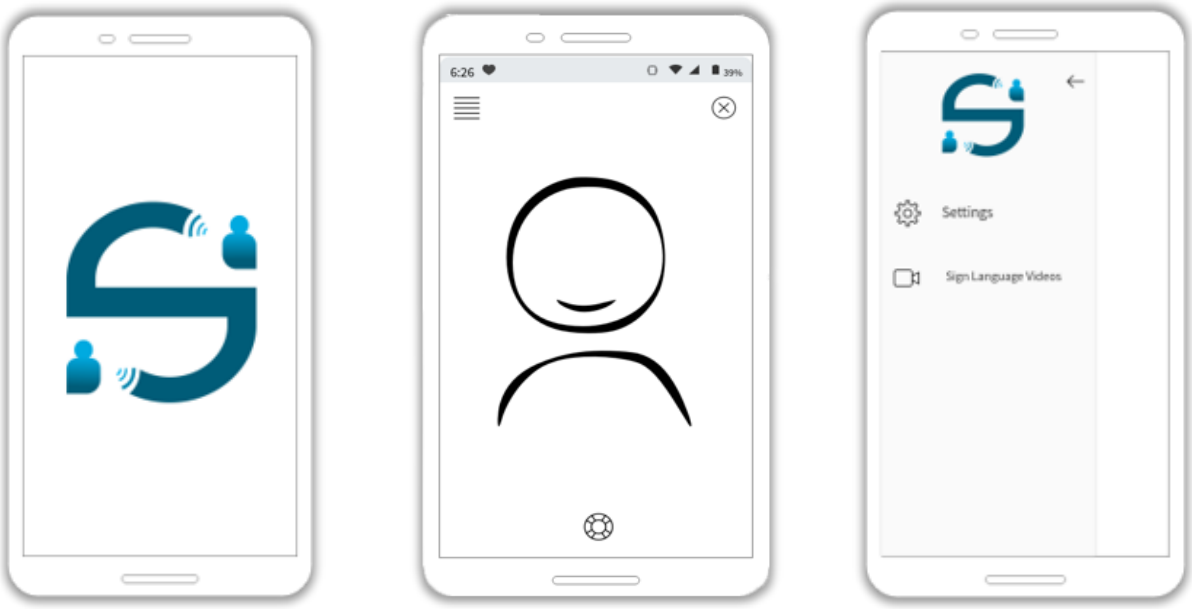

Fig. 2. User Interface consists of from left to right
a) Loading Screen.
b) Main Screen.
c) Side Menu.

SLARA's development process is divided into 2 stages:

- The first stage represented in Fig. 3 shows that the speaker's voice will be detected by the phone's mic and converted into words using GCSR. These words are mapped in a database with the corresponding SL movement that will be performed by SLARA and displayed into Vufine glasses within seconds.

  So, the normal person's words are translated to SL automatically without any need to learn SL to communicate with DM or an interpreter to communicate.

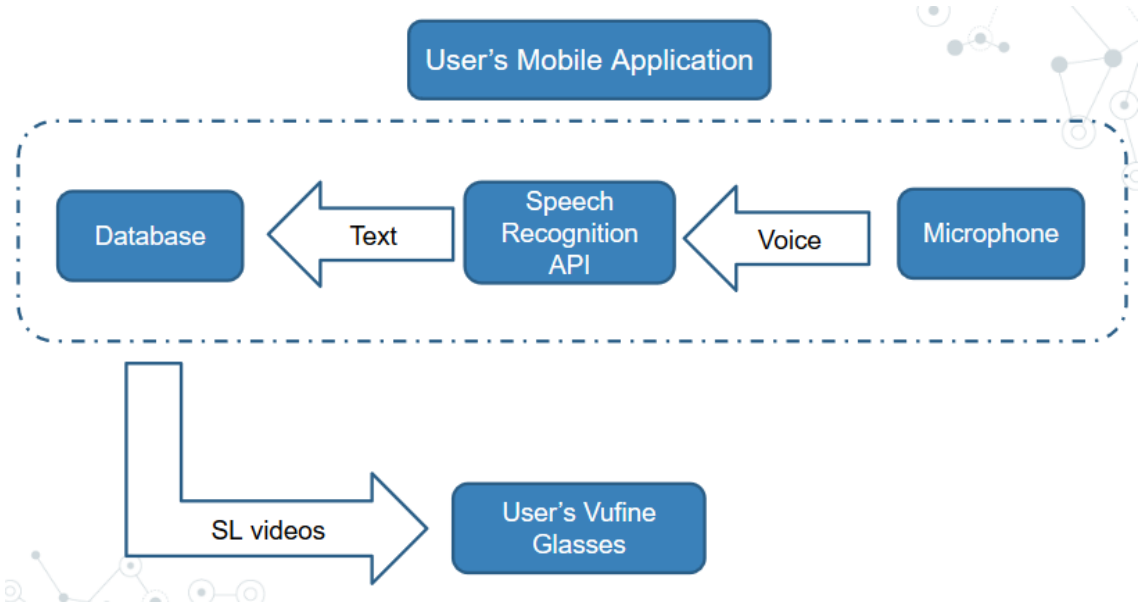

Fig. 3. First Stage

- Fig. 4 shows the second stage where the SL movements done by the user are detected by a camera using a tracking algorithm and Convolutional Neural Networks using Tensorflow machine learning library and OpenCV library. These movements will be mapped in the database to a corresponding text after processing to be converted into speech using text to speech plugin through the mobile phone's speaker.

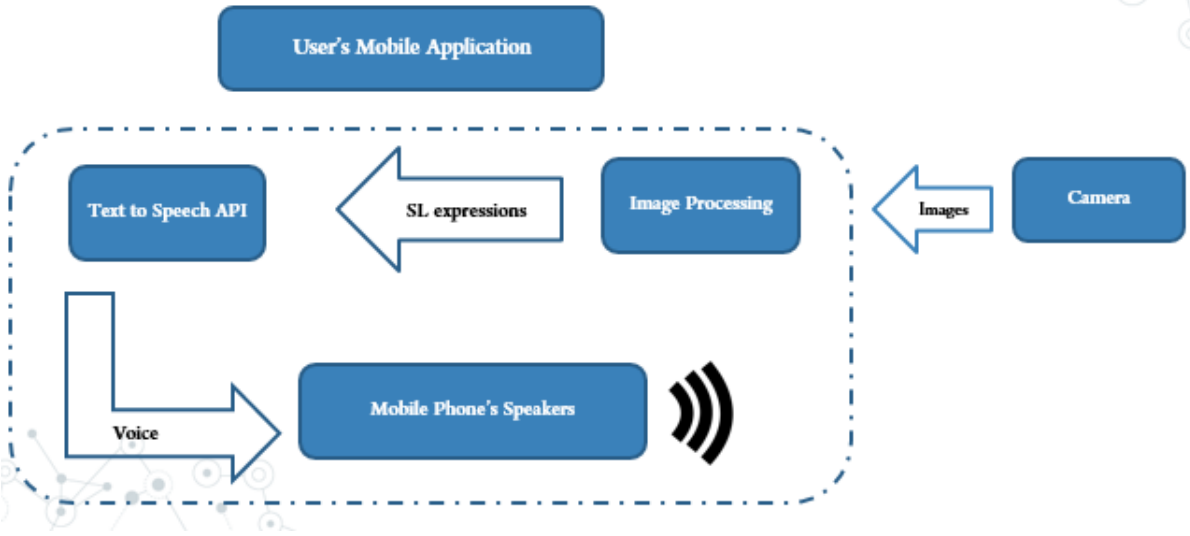

Fig. 4. Second Stage

For the enforcement of SL in our communities, Learning SL for communication became a necessity so an option is made in our SLARA which is retrieving the videos in order to learn SL with its corresponding text.

Cost

Component	**Price**
Vufine Wearable Display	**200 dollars**
Camera	**40 dollars**
GCSR	**20 dollars**
GCTS	**20 dollars**
Google APIs	**20 dollars (able to be increased monthly)**
Total cost	**300 dollars**

Results

SLARA project is concerned about the lives of DM people by offering a product that helps in communication with normal people, solves all the issues in the competing mobile applications and tries to increase the awareness of DM needs in our community.
A mobile application on the mobile phone which only an interface for the user to turn on the application and adjust the settings of it.

First, Using Vufine display to help the DM to see an avatar making a SL of the normal people's speech without holding a mobile phone that may cause a discomfort to the speaker and the DM (the listener).
Second, SL of the DM is interpreted as a speech to be listened to the normal people without any involution of any of them.
Third, even if the DM or the speaker is talked in another language, our project settings can be adjusted to support other languages.
Fourth, the camera is fixed above the Vufine display glasses with an angle to capture the hands of DM easily.
Fifth, all our processes are made in almost real time automatically.

Significance

We presented a mobile application with AR technology system for DM people to live a normal life without feeling disabled. The main features of this system are the comfort in use and automation of the whole conversation. The system allows normal people to communicate with the DM people without having any knowledge on SL or interpreter.

Acknowledgements and References:

1- After effects information available at: https://www.adobe.com/products/aftereffects.html
2- FFMPEG scriptable converter information available at: https://www.ffmpeg.org/
3- Filmora Wondershare information available at: https://filmora.wondershare.com/
4- Google cloud speech recognition information available at: https://cloud.google.com/speech-to-text/
5- Unity3D Game Engine information available at: https://unity3d.com/
6- Vufine information available at: https://store.vufine.com/products/vufine-wearable-display
7- Sign Language Recognition Application Systems for Deaf-Mute People: A Review Based on Input-Process-Output https://www.researchgate.net/publication/320402323_Sign_Language_Recognition_Application_Systems_for_Deaf-Mute_People_A_Review_Based_on_Input-Process-Output/
8- Recent Developments in Sign Language Recognition Systems https://ieeexplore.ieee.org/document/6707157/