SignAR: A Sign Language Translator Application With Augmented Reality

Noor-Un-Nissah B.N Soogund[1] and Minnu Helen Joseph[2]
[1]Faculty of Computing, Engineering & Technology, Asia Pacific University of Technology & Innovation, 57000 Kuala Lumpur, Malaysia

[1]noorsuho@mail.com, [2] minnu.helen@apu.edu.my

***Abstract*— Various studies have shown that deaf/non-hearing students face challenges in reading comprehension mostly because they cannot develop a fluent system of communication required to become proficient readers. This can affect their communication skills, leading to an overall depressed academic achievement and social seclusion. This is where SignAR comes into play. SignAR will help deaf children learn English as well as sign language and hearing people can also use it to learn sign language. So, basically, once a word has been captured by the phone camera, the corresponding sign animation will be augmented on the user's screen. By using SignAR, the deaf children will be able to learn English using Signs while hearing users will learn Signs using English. In this way, the communication barriers between the hearing and non-hearing will hopefully be greatly diminished.**

***Index Terms*— Augmented Reality, Deaf, Education**

1. Introduction

Reading is an important skill required to function successfully in today's society and it is a skill that is very much taken for granted by the hearing people. As a matter of fact, functional literacy is difficult to achieve for the deaf population. If there is one concern among teachers of deaf students, it is the challenge of reading comprehension. 90% of children with severe hearing impairment do not develop enough understanding of any language modality—whether it be oral communication or sign language to help in the process of understanding written language. Children born with a hearing impairment may not be able to develop a fluent system of communication required to become proficient readers (Magee, 2014).

Over the last 40 years, studies have shown that deaf children have considerably lower literacy skills, poorer reading comprehension, and overall depressed academic achievement when generally compared to their hearing peers, reducing the chance of being enrolled in postsecondary education institutions (Garberoglio, Cawthon, & Bond, 2014). Various surveys show that approximately 50% of deaf students in the US were reading below the fourth grade level at the time of their high school graduation with only 7–10% of deaf high school graduates reading at the seventh grade level or above. Moreover, reports have shown that the mathematical achievement of deaf students in many countries has been considerably poorer than that of their hearing peers. This means that deaf students have lots of trouble reading and understanding text, which is where the proposed solution will come and help them with the real time, instant translation of the signs.

To assist the deaf children in class, currently sign interpreters can be used, making the deaf student constantly dependent on the sign interpreter for the latter to sign the reading materials to them. When the sign interpreter is absent or there is no sign interpreter in the classroom, the deaf child will find it difficult and frustrating to follow the class and participate (Mahwish Safder, 2012).

Another problem that can affect the education of a deaf child is that teachers do not know how to use sign language. As a result, the deaf child cannot follow instructions and write down notes. (Mpofu, 2013)

From the above statement, it can be concluded that children with a hearing impairment have reading difficulties and this can affect their communication skills. When their communication skills are affected, their future job prospects are ruined, making them forever dependent on people, whether it is a sign interpreter or their parents. Their case is even made worse by the majority of hearing people who don't know sign language. It is for these above reasons that the researcher has decided to develop the Sign Language Translator which will be used by both hearing and non-hearing people. By scanning a text, a virtual sign translator will be shown on the user's screen, making the signs. In this way, the child's reading skills can improve as the English words/ text are mapped to the sign language.

Due to the existing problems the deaf students are facing, the SignAR will help these deaf children overcome these problems. The proposed solution by the developer is an Augmented reality application that will provide 'invisible' sign language mapping of every word in a text. In other words, the AR application will replace the human interpreter. This AR application will help deaf or hearing-impaired students immensely as it will enable them to improve their signing by learning more words through reading and they can learn at their own pace. This app is also helpful for the hearing to learn how to sign simply by placing the screen onto text or an object and thus, through the use of this AR application, the communication problems between the hearing and the deaf will be diminished. It will be easier for deaf/hearing-impaired people to read menus at restaurants easily, easy for them to follow a class or read a comprehension.

2. Literature Review

Augmented reality (AR) is a branch of computer science research that merges real world and digital data where the aim is to combine the physical reality together with computer generated graphics. As a technology, augmented reality is now on the top of the "technology hype curve". New augmented reality applications are being introduced all the time. (Siltanen, 2012).

Below is a definition of Augmented Reality that can be found in a survey by Azuma (Azuma, 1997):
This survey defines AR as systems that have the following three characteristics:

1. Merges real and virtual, meaning that animations or information are displayed on the same display as the user sees the real world on or through, in order to improve whatever the user sees and reads out from real world objects, that is the deaf child will be able to see the signs done by a virtual interpreter on his mobile device
2. Interactive in real time, that is, the user can interact with the information displayed in the AR system, for example the user might look at an object and is provided information like the sign language about it and by clicking on this information the user can save the translation
3. Registered in 3-D means that virtual information is shown and aligned with the real-world object, that is the signs are overlaid next to the image or text

An augmented reality system is built up of a camera, a computational unit and a display. The camera captures an image, and then the system augments virtual objects on top of the image and displays the result. This can be achieved by using an SDK like Vuforia platform which tracks and registers the objects of the real world and matches them with the objects of the virtual world. The system captures an image of the environment, detects the marker and calculates the location and orientation of the camera, and then augments a virtual object on top of the image and displays it on the screen.

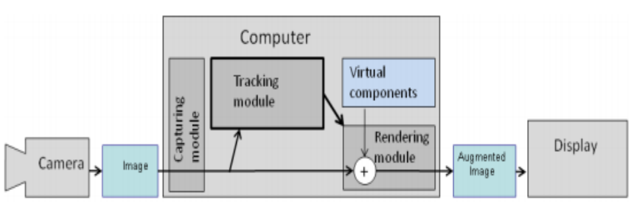

Figure 1 showing a simple augmented reality system

Figure 1.0 shows a flowchart for a simple augmented reality system. The capturing module captures the image from the camera. The tracking module calculates the correct location and orientation for virtual overlay. The original image and the virtual components are merged by the rendering module using the calculated pose and then renders the augmented image on the display.

In order to achieve the above, there are various approaches used by the AR community. A number of these approaches belong to two categories: pose computation and object recognition. A while ago, the algorithms for object detection and pose tracking have been introduced in AR applications. In these systems, various objects are registered first on a database. When one of the registered objects is scanned by camera, it is recognized, tracked and localized to superimpose its corresponding virtual object on it. Object recognition is used by many in the AR community to provide information about the detectable objects captured by the user. However, this technique will work only if the user is viewing specific known objects; it does not support augmented reality when there are no registered objects in the scene. (Z. Kalal, 2012)

Object detection works through Natural feature tracking which needs better image quality and more computational resources, and has only recently become popular. The wide availability of mobile devices with built-in cameras makes this the preferred tracking hardware for mobile AR. (Ufkes, 2013)

Natural Feature Tracking is an image-based tracking technique that recognizes and tracks the features that are naturally found in the image itself. These could be corners, edges, blobs without using specifically designed ID markers. Below shows a target image uploaded for SignAR on the Vuforia database. Vuforia provides an option to see the "natural features" along with the augmentable scale of the image.

Figure 2 showing the natural features (the yellow crosses) detected

For SignAR, an approach of that type will be used. The words to be recognized are already registered into the Vuforia database. The same thing goes with the images that the developer wants the SignAR application to recognize and detect. The images are uploaded on the Vuforia target database. The SDK of that database is then downloaded and imported to

Unity 3D where the developer can superimpose the respective animations on each word or image from the database.

As mentioned above in this literature review, Vuforia SDK incorporates both object recognition and Text recognition. In this section the researcher will talk about how the text recognition actually takes place in Augmented Reality.

Optical Character Recognition (OCR) is a technology that detects and recognizes printed text inside an image and converts it into digital format (P.C.Pop, 2013). The group of pixels that represent a letter are compared to the shape of the actual character so that the equivalent can be returned. In contrast, there is also Intelligent Character Recognition (ICR), an engine that enhances the character recognition by reading handwritten text, with a neural network as a self-learning system. However, having both speed and accuracy can be a true challenge in OCR. When facing the real world, there are many issues to be considered, such as low resolution, picture distortion and rotation, heavy noise or damaged data. Dealing with accuracy can require heavy programming, whereas if the goal is speed, the results can be less precise. (Z. Kalal, 2012)

Both Natural Feature Detection and Text Recognition are incorporated into the Vuforia SDK which uses Image Targets to detect the images. Image Targets do not require special black and white regions or codes to be detected. The SDK recognizes and tracks the features that are naturally found in the image itself by comparing these natural features against a known target source in the database. Once the marker is detected, the SDK will track the image as long as it is at least partially in the camera's field of view. In the same way, the Vuforia Text recognition will be used to detect text and display an animated avatar on the screen.

3. Research methodology and discussion

The Scrum methodology was used to develop SignAR mainly because it is the methodology most recommended for mobile applications that have to develop over a short period of time. This reflects the case of the SLTA app which needs to be built over 6 months or so. Also, in this case, the student's supervisor is acting like the Scrum master who will monitor the progress and conduct consultations with the student, it is almost like the student who is also the developer will be conducting a sprint where she will come up to discuss about the possible increments about the system with the supervisor also known as the Scrum master.

The Scrum methodology aims at being flexible all along the project life. It offers control mechanisms for planning a release, and then for handling the project as it progresses. This helps the developer to change the project and its deliverables at any time, delivering therefore the most appropriate release. The Scrum methodology gives freedom to the developer, so that she can focus on developing innovative solutions during the project, taking into consideration the learning curve and changes in the environment. This is something to consider given that the developer will need to learn many new techniques such as animation on Unity and creating AR apps. So sufficient time has to be given to the developer while taking into account the learning curve. This is another reason why Scrum is considered an appropriate methodology to follow. (Schwaber, 2013)

3.1 Requirements Gathering

In the requirements gathering phase of the Scrum, an online questionnaire was used to gather the data and requirements of SignAR because it is really easy to gather the information required by the developer using this method. Finding deaf or hard of hearing people can be hard, hence using the questionnaire the researcher will ask Internet users who are deaf/ hard of hearing to fill out the questionnaires. This has been made even simpler because of Social Networking sites like Facebook that has circles or groups for Deaf/ hard of hearing people making it even easier for the researcher to get in touch with them.

The demographics of the respondents are listed below:

- Gender: 56.6 % Male; 44.4% Female
- Age range: 50% were above 20 years;16.7% between 10 and 15; 33.3% between 15 and 20
- Occupation: 72.2% student; 16.7% jobseeker

The main objective of the questionnaire was to gain an understanding of the requirements of the deaf/ hearing-impaired people for SignAR and whether they would like to have such a system as an educational tool. It also aimed at finding out the problems they were facing while studying and the communication barriers they face daily.

To get more knowledge on the challenges faced by non-hearing people and also to increase usability of SignAR two interviews were conducted with them. This helped the developer see the problems faced by them in a clearer picture and the majority of the respondents agreed that scanning a text and getting the animated signs on the screen is much better than having to type them.

After the requirements were gathered, the developer started designing the system and started coding at the same time. The developer would conduct sprints with the supervisor where a little increment was shown each time.

After the implementation, User Acceptance testing was carried out as well as Unit testing for each functionality. A questionnaire was given out to three non-hearing people and their feedback was recorded while they used SignAR.

4. Results and Discussion

From the data gathering techniques used, around 70% of the respondents agreed that in class, it is difficult to follow the lecturer who doesn't know Sign language. This shows that SignAR can be used by hearing people to learn sign language. When asked whether the non-hearing have difficulty in understanding some words and how to read them, 77.7% agreed to that. Again, this challenge can be overcome as the animated signs will be superimposed on the screen onto each word.

Another important response was that 57.9% of the respondents strongly agree that they don't know how to sign a word, hence further strengthening the credibility of SignAR which will help learning new signs by scanning the word.

The interview response helped to give more insight into the problems faced by deaf/non-hearing students, with some agreeing that making friends or following a class is really difficult without a sign interpreter during the class. The response also showed that the interviewee agreed that scanning a word is better than typing it to get the translation on the screen as this feature can also be used by disabled people who can't use their fingers.

On the whole the majority of the respondents felt that more should be done to accommodate deaf/hearing impaired students in the classroom and that more people should learn sign language. Hence SignAR would be really useful and helpful.

Figure 3 below shows the first scene the user will see when SignAR is launched. The purpose of this scene was to introduce the virtual sign Interpreter, Harry, so that users can feel like it is a friend with whom they can connect. The sign interpreter is actually animated and waving "Hi" on the screen. This is the 3D model the developer used to create the other animations for the words in Maya Autodesk.

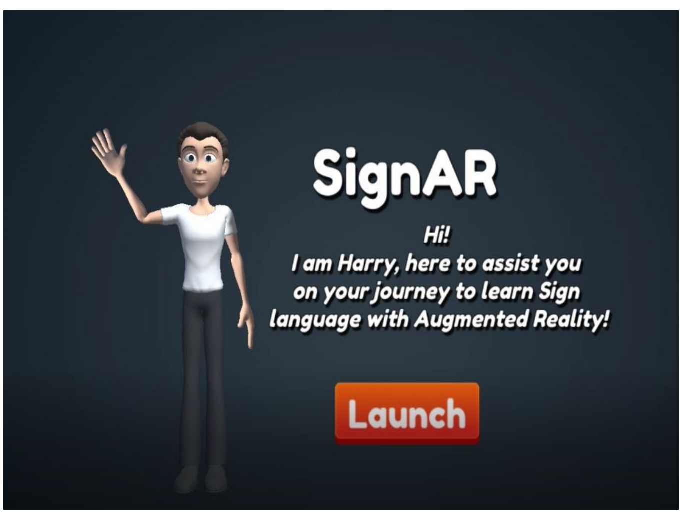

Figure 3 shows the 3D model sign interpreter

Figure 4 shows the words that are being recognized with the virtual sign Interpreter displaying the animations through the smartphone's screen:

Figure 4 shows the sign interpreter making signs based on the words recognized

The animations for "The", "Who", "swims" and "across" are being displayed. The user can also scale up and down the animations to better see the animation and the hand gestures.

Figure 5 below illustrates the image recognition feature of SignAR. For the purpose of SignAR, a specific book was considered and pictures of animals were taken and registered onto the Vuforia database. On detection of each animal, their respective signs will show up:

Figure 5 shows the sign interpreter making the "cow" sign because the cow picture was detected

After the system was developed fully, the developer made 3 non-hearing users use the system. The response was amazing and highly motivating for the developer. They all loved the idea and agreed that SignAR will be an indispensable tool in deaf education. Some feedback helped the developer improve the user interface like the positions of the button. Other feedback gave the developer more room for future enhancement like adding a button to translate the whole sentence.

5. Conclusions

Based on the user acceptance testing, the developer can say with confidence that SignAR is working as expected according to the **core** requirements which are:

- Scanning an image to get the animation sign
- Scanning a text to get the animation sign

All of the above features are working perfectly. Since the developer had additional time, she has also implemented some additional features like:

- Sign up to Firebase database
- Login to Firebase database
- Allowing the user to scale up and down, move the animations away from the word.
- Sending feedback to the developer and seeing it live on the app

All of these features make SignAR a better and much more user-friendly app when compared to the apps which are Mimix, ProDeaf and ASL Translator. In the case of SignAR, the user doesn't need to type the words and keep pressing on replay, which can be annoying and tiring. Instead, the user just needs to capture a text or an image to get the signs. In addition to that, the sign animation is on loop, so there is no need to keep pressing the replay button. This is perhaps the biggest advantage of SignAR, implying that even people with diseases like arthritis (a disease that make the joints in fingers swell and get stiff) can use SignAR to communicate or learn sign language. Since the signs are being done by an animated character, kids will be more fascinated and interested to learn. This is reinforced by the fact that the user can interact with the animated character while it appears in the real world, making learning more fun! It almost gives a feeling of holograms as compared to the other traditional apps mentioned above that appear in the screen only as a 2D image. The developer therefore concludes that SignAR is a beginning step towards breaking communication barriers between the deaf and hearing people and that nothing is impossible if one really works towards that goal. SignAR seemed impossible to build at the beginning because there has never been any AR application like that, and yet the developer has managed to do it through perseverance and patience!

6. References

Azuma, R. T., 1997. A Survey of Augmented Reality. *Teleoperators and Virtual Environments.*

Ćuković, S., 2015. Marker Based vs. Natural Feature Tracking Augmented Reality Visualization of the 3D Foot Phantom. *Proceedings of the International Conference on Electrical and Bio-medical Engineering, Clean Energy and Green Computing.*

Magee, P., 2014. CHALLENGES WITH LITERACY DEVELOPMENT IN CHILDREN WHO ARE DEAF OR HARD OF HEARING. *Southern Illinois University Carbondale* .

Mahwish Safder, M. M. S. A. G. F. M. M., 2012. Problems Faced by Students with Hearing Impairment in Inclusive Education at the University Level. *Journal of Research and Reflections in Education.*

Mpofu, M. J., 2013. Challenges faced by Hearing Impaired pupils in learning: A case study of King George VI Memorial School. *IOSR Journal of Research & Method in Education.*

P.C.Pop, H. O., 2013. Optical character recognition in real environments using neuralnetworks and k-nearest neighbor. *Springer Science+Business media New york.*

Pankratz, 2013. An interactive augmented reality platform for CAD education. *Proceedings of 35th Conference on Production Engineering.*

Schwaber, K., 2013. The Scrum Guide.

Siltanen, S., 2012. *Theory and applications of marker-based augmented reality.* s.l.:s.n.

Ufkes, A., 2013. A Markerless Augmented Reality System for Mobile Devices.

Z. Kalal, K. M. a. J. M., 2012. Tracking-Learning-Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence.*