International Journal of Information Science and Education.
ISSN 2231-1262 Volume 2, Number 1 (2012) pp. 13-20

http://www. ripublication.com

An Avatar Based Translation System from Arabic Speech to Arabic Sign Language for Deaf People

S.M. Halawani and Zaitun A.B.

Faculty of Computer and Information Technology,
King Abdul Aziz University at Rabigh, Kingdom of Saudi Arabia
E-mail: dr.halawani@gmail.com, zaitunab@gmail.com

Abstract

The advancement of Information and Communication Technology has impacted all aspects of the human life. It has changed the way we work, travel, conduct business, study and communicate. For the Deaf community, we try to use ICT to improve their quality of life by developing systems that can help them communicate better with the rest of the world and amongst themselves. In this paper we present to the reader our research on the development of an Avatar based translation system from Arabic Speech to the Arabic Sign language for the Deaf people. The paper starts with an introduction of the problem of the Deaf community in the Arabic speaking community and a brief review of similar and related research. The next section describes the design consideration of the proposed system. The system will be composed of a database of captured 3D motions of Arabic sign language. The sign language motion will be recorded using data gloves. A Graphical translation of the digitized sign language will be re-animated using standard. The emphasis is to produce the original gesture as close as possible. Common spoken words are directly translated into respective semantic using Sphinx-4 Speech Recognition Engine without first being translated into text. Using the same Sphinx-4 Engine, the semantic of the spoken Arabic language will then be translated into Arabic sign language. With minimum supervision, the automation of the translation for converting standard Arabic speech into virtual Arabic Sign Language will be useful to the hearing impaired Arabic speaking community. Finally we describe future enhancements and other research projects that can be derived from this research.

Keywords: Arabic Sign Language, Avatar, Speech Recognition, Deaf Community

Introduction

ICT has been used to improve the quality of life for the disabled community. Significant innovations and inventions have been introduced in helping people with learning disabilities overcome their problems. We should also explore how we can utilize ICT in helping the deaf community in overcoming their communication problem. The world population has just touched 7 Billion in 2012. According to the World Federation of the Deaf [1], the total number of deaf people worldwide is around 70 million and the number of deaf people in Saudi Arabia alone is 1,103,284 [2]. Therefore working on a research project to develop a system to translate an Arabic speech to an avatar, will have a significant impact to the deaf community in the Arab world.

Deaf people communicate using sign language, which is a communication system by gestures as opposed to written or spoken language. There are 25 countries where Arabic is an official language. In some countries Arabic is spoken by a minority of the people. Some sources put the number at 22-26 countries. Even though the Arabic language is diglossic, fortunately, the Arabic sign language is not. The Jordanians, Egyptians, Moroccans, Libyans, Iraqis and Palestinians, for example, all speak Arabic. However, each country has its own dialect. In other words, there are basically two types of Arabic; standard and colloquial. Fortunately, they all use the same alphabets and therefore the Arabic sign language (ArSL) is the same. This feature helps a lot in our research project. The Arab deaf communities are almost closed ones. Interaction between the deaf community and a hearing one is minimal and is basically concentrated around families with deaf members, relatives of the deaf, and sometimes play friends and professionals. ArSLs are basically manual languages that involve the three recognized elements: configuration of hands (hand shape), placement/space (position of hand in relation to body), and movement (directions and contacts within space). In addition to these manual shapes, ArSLs make use of other non manual features, like those of the face, mouth, and tongue. [3].

There is a need therefore for a translation system that can convert an Arabic speech to an ArSL and vice versa, so that the deaf community can communicate better with the normal people. In the next section we will describe some related works that have been contributed by other researchers.

Related Work

In 1994, at the Rehab R&D Center of Stanford University, USA, “Ralph” was invented. "Ralph" simply stands for Robotic Alphabets. It is actually a fourth generation computer controlled electromechanical fingerspelling hand. "Ralph" provides a deaf-blind person some independence in communication. The device translates key presses from a keyboard into movements of the fingers of a mechanical hand. These movements are felt by a deaf-blind user and interpreted as the fingerspelling equivalents of the letters comprising a message. These finger movements enable the user to receive finger spelled messages from the mechanical hand in response to person-to-person communication as well as gain access to sources of computer-based information. With a fingerspelling hand, a deaf-blind individual need not rely on a human interpreter for all communication [4].

In 2002, Ryan Patterson invented the CyberGloves, which incorporates virtual reality sensors to capture both isolated and continuous hand signs. The glove is fitted with sensors that can detect or track the signs that an individual makes. Special software is then used to translate the hand movements of the sign language alphabet and wirelessly transmitted the data to a portable device, from where it will be displayed on a computer screen [5].

CopyCat is a game which is designed to help deaf children develop working memory and language skills while playing. It uses camera-based technology and sensors to capture and collect gesture data for the American Sign Language (ASL) recognition system. The system uses a video camera and wrist mounted accelerometers as the primary sensors. A user will wear a colored gloves and any gesture made will be captured by the camera. The system then shows a short video with a signer demonstrating the correct ASL phrase. The user can then mimic these gestures. Gesture-based interaction expands the possibilities for deaf educational technology by allowing signing children to interact with the computer using their gesture-based language [6].

The ICICLE system (Interactive Computer Identification and Correction of Language Errors) [7] is an intelligent tutoring system under development in the University of Delaware, USA. The main objective of ICICLE is to utilize natural language processing and generation to tutor deaf students on their written English. A student submits a piece of writing and the system will review it. The system will then perform a syntactic analysis, determines its errors and provide a tutorial feedback to the student. The feedback will make the student aware of the nature of the errors found and provide the necessary information to correct the errors. When the student has corrected the errors, the writing can be resubmitted to the system and the cycle is repeated. This is indeed a novel contribution to the deaf community.

Another example of a computerized system to help the deaf community is THETOS, which is actually a spoken language to sign language translation system for the Polish sign language. The system uses natural language processing techniques [8].

Another interesting project to aid deaf people is Tessa [9]. It is an experimental system to help deaf people with their transactions with a clerk at a post office, in the UK. It uses avatar and the British Sign Language (BSL). Tessa uses a speech recognition component which translates the speech from the clerk and then generates the appropriate sequence of signs in BSL using a specially developed avatar.

We have so far described systems for the deaf people using ASL, BSL and the Polish language. Halawani [10] has ventured further by developing an automated Arabic Sign language translation system. The output of the translation can be displayed on to wireless devices using Wireless Application Protocol (WAP).

We have described some examples of automated system to help the deaf community to communicate better and improve their quality of life. These systems have used various techniques such as Robotics, Virtual Reality, Natural Language processing, camera based technology and sensors. These systems are dedicated mostly for ASL and BSL and other European Language. Halawani's work uses ArSL and involves mobile devices. The objective of our research project is to develop an automated system to capture a speech in Arabic and translate it to the ArSL using

Avatars. In the following section, we will describe the architecture of the system, which we call ABTS for ArSL (Avatar Based Translation System for Arabic Sign Language).

Design and Development of the Proposed System

The system will be made up of two parts. The first part receives the input in the form of a speech which is in Arabic. The speech is captured through a speech recognition system as an analogue signal. Then it is digitized and translated to Arabic sign language. The second part of the system converts the Arabic sign language to avatars which are displayed on the computer screen for the user to see. The system will refer to two sets of databases. The first database, which we call the Sign and Speech database will contain all the images of the alphabets and words of the ArSL and the recorded Arabic words. While the second database called the Avatar database will contain the equivalent avatar for each alphabet and words in the ArSL. Figure 1 below describes the overall architecture of the ABTS for ArSL.

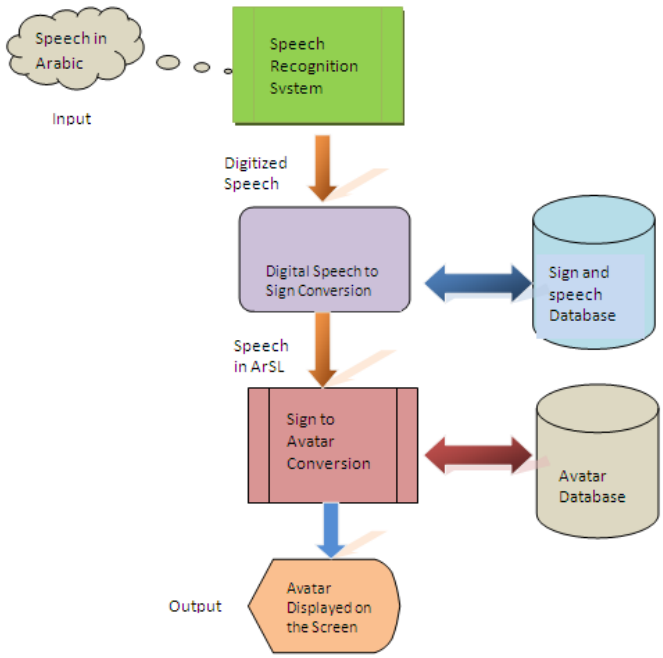

Figure 1: Overall Architecture of the ABTS for ArSL

The Speech Recognition Engine

Speech recognition is the process of converting an acoustic signal, captured by microphone or a telephone, to a set of words. The recognized words can be the final results, as for applications such as commands and control, data entry and document preparation. They can also serve as the input to further linguistic processing in order to achieve speech understanding, a subject covered in section. As we know, speech

recognition performs their task similar with human brain. Start from phoneme, syllable, word and then sentence which is an input for speech recognition system.

Many researchers have experimented on how to decrease the error and also any disruption while doing the recognition. For the speech recognition component the process is summarised in Figure 2 below. We have chosen to use Sphinx-4 which is a state-of-the-art speech recognition system written entirely in the Java™ programming language. Sphinx-4 is a very flexible system capable of performing many different types of recognition tasks [11].

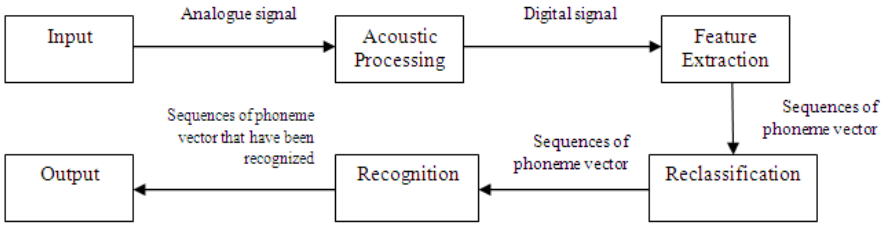

Figure 2:-The Speech Recognition Process

The Sign Database

The ArSL is almost similar to ASL and BSL [3]. It is basically made up of three elements: configuration of hands (hand shape), placement of hand / space (position of hand in relation to body) and movement (directions and contact within space). Apart from these manual shapes, ArSL also makes use of other non manual features like those of the face, mouth and tongue. For the sign database we store all the avatars of the 29 Arabic alphabets as depicted in Figure 3 below.

Figure 3: Signs of the 29 Arabic Alphabets

The Avatar Database

In the context of our research project, an Avatar is simply the graphical representation of the word that we want translate to ArSL. Therefore the Avatar database will be developed to contain as many symbols or graphical icons as possible to represent Arabic alphabets and words. For example; we can use the 3D images and videos in Figure 4 to represent Arabic alphabets.

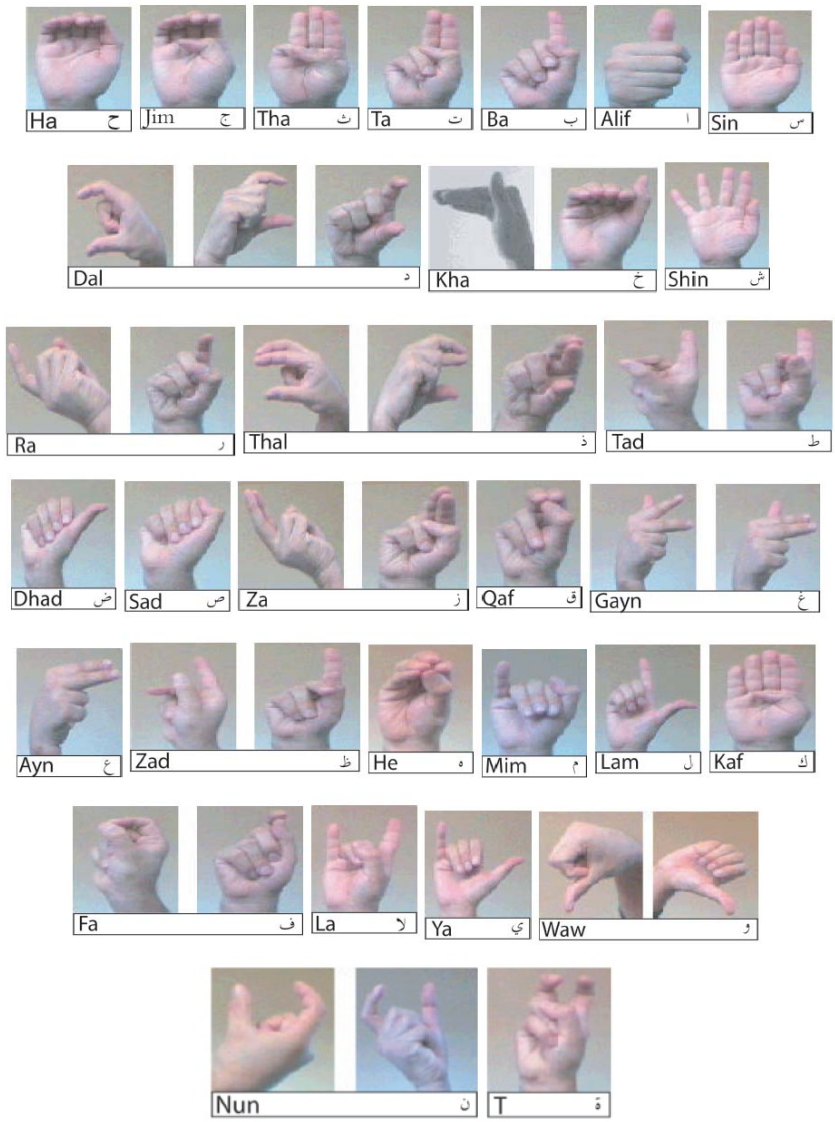

Figure 4-Gesture of Arabic Sign Language

We can also use the avatars in Figure 5 and 6 to represent Arabic words. The more avatars we have the more efficient will the system be.

Figure 5: An Avatar Representing the word "Rectangle"

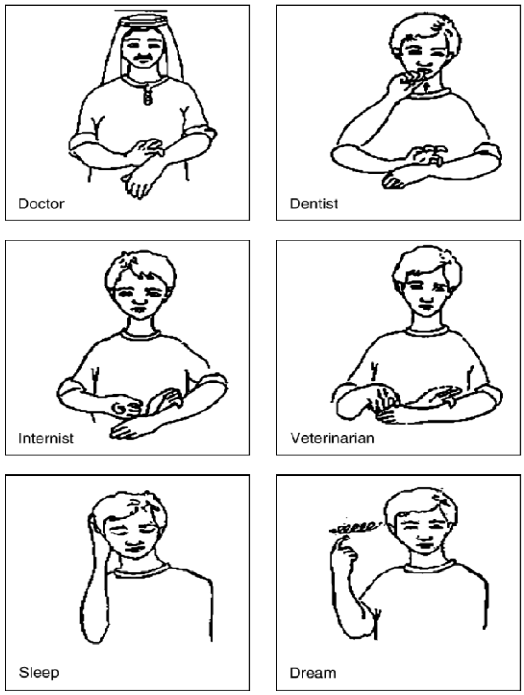

Figure 6:-Avatars Representing Words.

Conclusion and Future Work

Sign language recognition and translation is still an active area of research and will continue to evolve. In this paper we have described the proposed ABST for ArSL system to translate a speech in Arabic directly to avatars. For this project, the biggest challenge will be developing the speech recognition system. The accuracy of capturing Arabic speech and deciphering it correctly is vital for the success of the following processes. The building of the sign database can be time consuming but nevertheless, it is achievable with the right resources. The same can be said for the populating of the Avatar database with avatars. The next biggest challenge will be to select and use the best technology available. At the moment we are presented with many possible technologies such as; Robotics, Virtual Reality, Computer Vision, Neural Networks, Hidden Markov Models, 3D Animation and Natural Language Processing. They can be used to complement each other and this have yet to be explored in Sign language translation systems research. For future expansion, we can consider catering for the diglossic nature of the Arabic language and also a web-based version of ABST for ArSL, so that users can use it from anywhere at any time. Another enhancement will be to implement the system on mobile devices.

Acknowledgement

We wish to thank to King Abdul Aziz University for awarding the research grant and

the lecturers of Faculty of Computing and Information Technology for their continuous support and help. We also wish to thank to our fellow researchers who have assisted us throughout the project.

References

[1] World Federation of the Deaf, available at, http://www.wfdeaf.org/ last access date: 22 January 2012.

[2] Ethnologue.com, available at; http://www.ethnologue.com/show_country.asp?name=SA last access date: 22 January 2012.

[3] M. A. Abdel-Fattah, Arabic Sign Language: A Perspective, Journal of Deaf Studies and Deaf Education vol. 10 no. 2, 2005. Oxford University Press.

[4] Watters S., Miller T., Balachandran P., Schuler W., and Voyles R., Exploiting a sensed environment to improve human-agent communication, AAMAS '05 Proceedings of the Fourth International Joint Conference on Autonomous Agents And Multiagent Systems. ACM New York, USA ©2005

[5] Huenerfauth M. and Lu P., Accurate and Accessible Motion-Capture Glove Calibration for Sign Language Data Collection, ACM Transactions on Accessible Computing (TACCESS), Volume 3 Issue 1, September 2010, New York, NY, USA

[6] Brashear H., Improving the efficacy of automated sign language practice tools, ACM SIGACCESS Accessibility and Computing - ASSETS 2007 doctoral consortium, Issue 89, September 2007, ACM New York, NY, USA

[7] Michaud L. N. and McCoy K. F., Modeling user language proficiency in a writing tutor for deaf learners of English, ASSESSEVALNLP '99 Proceedings of a Symposium on Computer Mediated Language Assessment and Evaluation in Natural Language Processing, Association for Computational Linguistics Stroudsburg, PA, USA ©1999.

[8] Kennaway J. R., Glauert J. R. W. and Zwitserlood I., Providing signed content on the Internet by synthesized animation, ACM Transactions on Computer-Human Interaction (TOCHI), Volume 14 Issue 3, September 2007, ACM New York, NY, USA

[9] Cox S. , Lincoln M., Tryggvason J., Nakisa M., Wells M., Tutt M., Abbott S., Tessa, a system to aid communication with deaf people, Assets '02 Proceedings of the fifth international ACM conference on Assistive technologies, ACM New York, NY, USA ©2002

[10] S. Halawani, Arabic Sign Language Translation System on Mobile Devices, IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.1, January 2008.

[11] Sphinx-4, available at http://cmusphinx.sourceforge.net/sphinx4/#what_is_sphinx4, last access date: 18 January 2012.