

# Meet the First Qatari Sign Language Avatar: A 3D Realistic Virtual Conversational Agent

Achraf Othman

IEEE Senior Member

achraf.othman@ieee.org

## **Abstract-**

When it comes to inventions and technological assistance for the hearing impaired, science has come a long way, there is absolutely no doubt about that. However, we live in a day and age where humans, by default, are genetically engineered to want more and to crave more, which is completely acceptable. With standards of living going up with each passing sense, it makes sense that individuals with any kind of impairments demand more than just hearing aids. On September 28, 2021, Mada Center launched the first of its kind in the world, a 3D realistic virtual conversational agent for Qatari Sign Language with the aim to enhance ICT Accessibility in the Arab region and beyond. This article will present a brief overview of the avatar technology for sign language and an introduction about “Bu Hamad” the Qatari Sign Language Interpreter.

**Keywords:** Computational Sign Language Processing, Realistic Avatar, Qatari Sign Language  
*Published on September 30, 2021.*

## **Introduction**

Thankfully today, with the immense leap forward of Artificial Intelligence, more research is being undertaken to come up with better solutions. One such invention is the Avatar technology or the 3D interpretation technology. On September 28, 2021, Mada Center launched the first of its kind in the world, a 3D realistic virtual conversational agent for Qatari Sign Language with the aim to enhance ICT Accessibility in the Arab region and beyond. This article will present a brief overview of the avatar technology for sign language and an introduction about “Bu Hamad” the Qatari Sign Language Interpreter. The innovation proposed by Mada Center is considered as a cutting-edge technology because it is based on the latest advances on Artificial Intelligence and Big Data. The data used to interpret a written Arabic Text to Qatari Sign Language was captured from hundreds of wearables sensors. The reason behind all of this is to be sure that the avatar “Bu Hamad” (Fig. 1) is taking in consideration all components of Sign Language. In the work of (Othman et al., 2019), a list of components of Sign Language is defined. This project was supported by the Mada Innovation Program (MIP) for the period 2019-2021 (Al Thani et al., 2019).

## **What is the Avatar technology?**

As the name suggests, the Avatar technology (Kipp et al., 2011) is basically the 3D translation of any closed captions using virtual conversational agent. The idea is that there is a certain interface, artificial intelligence or otherwise, which can understand the text, will automatically generate the



written text into sign language for the deaf individual to understand (Othman et al., 2019). This is called machine translation for sign language (Abeillé et al., 1991). In literature, we may find several techniques for machine translation. For example, there several works related to statistical machine translation (Othman et Jemni, 2011) and example-based machine translation (Somers, 1999). This interpretation is constructed based on a dictionary of words and signs that is already fed to the device (Jemni et al., 2013) (Othman et al., 2012). People can keep adding things or words to this dictionary, but only when their words are approved by a set of learned linguists and human interpreters (Tmar et al., 2013).



Fig 1. “Bu Hamad” the first Qatari Virtual 3D Interpreter for Qatari Sign Language by Mada

### **Why would anyone need the Avatar technology?**

You see, it is not exactly fair to expect every single hearing-impaired individual to learn sign language as well as all the languages that comes along with functioning in an everyday environment and living an average life. Every sign language has its fair share of grammar, and it gets difficult to translate that grammar into visuals. Then, there is the issue of basic differences between the kinds of sign languages in various other languages. For instance, the American Sign Language (ASL) is different from the German Sign Language (DGS). Therefore, a hearing-impaired German citizen cannot completely communicate with a hearing-impaired Indian citizen. Another issue is the lack of understanding of sign language on the part of the listener. In an everyday situation, it cannot and is not possible to expect that a completely abled person will know and use sign language, which is a problem since signing must be a two-way form of communication. Therefore, the deaf individual will be left at a loss.

At this point, the 3D technology would prove to be tremendously useful, because, as the idea goes, the interface would have a whole lot of sign language gestures fed to it. This would help both individuals communicate without the barrier of language. At the same time, it would also help individuals who do not know sign language communicate with deaf individuals.

### **Singing through Virtual Reality**

One of the ways in which 3D is being widely used to make sign language easier and more universal, is through capture motion. The basic idea is that an individual wears motion capture gloves and keeps signing with these, and then, the computer, equipped with motion tracks makes



a record of those movements and directly translates it into text or visual language. The very idea of this saw light back in the year 2002, when a teenager by the name of Ryan Patterson developed such a glove that understood and comprehended the signs made by the person wearing the gloves and proceeded to generate a written text and sent it directly to a portable device.

Of course, over the years, more and more complex and multitasking ways of such 3d capture have emerged to take the pace of this kind of glove. For instance, there are the neural networks, which take this to the next step by not only capturing the hand motion, but also detecting the slight changes or aberrations in them from person to person, and even detect or develop a significant pattern between such hand gestures.

### **Why the need for 3D translations of sign language?**

It is understandable that people may ask the need to translate closed captions into Sign Language. Especially when it comes to written texts, deaf people can read, so why this technology? You see, a deaf person's first language is not their native language, but their native sign language, and native sign languages differ from one country to another. For deaf individuals, who primarily learn their sign language, learning a written language is harder. For instance, for an average deaf American, learning English is harder than learning that American Sign Language. This is the reason why a lot of deaf individuals have issues reading and writing.

Thus, a lot of websites, to make written materials more readily available, use videos where someone automatically signs the written text. However, the major problem is that such videos need to be completely edited from the beginning, whenever the written text is edited. This costs time and money. This is where 3D signing comes in. The visually captured motions of sign language are first fed to the 3D avatar. This is then presented as sign language with motion blending and converted to a language understandable by the machine (Othman and Jemni, 2019).

### **What is the state of the 3D Sign language translation today?**

Unfortunately, although sign language is not universal, not a lot of progress has been made when it comes to the Avatar technology, and more funding is required for research. The major cause of this, is of course, the lack of a common sign language. Since languages keep on changing from region to region, it becomes extremely difficult to feed the signs of one word from every single language into the dictionary of the interface, not to mention that creating such a huge dictionary would cost a lot of money and would basically render the machine inaccessible to almost most of the people. Proposals are being set forth about the setting up of a common community which decides on a universal sign language, which, incorporate the signs and symbols from the major or the most spoken sign language. Moreover, several topics are not yet elaborated in depth to understand the nature of Sign Language such as the recognition of prosodic pauses in Sign Language (Lagha et Othman, 2019).

### **References**

- Abeillé, Anne, Yves Schabes, and Aravind K Joshi. 1991. "Using Lexicalized Tags for Machine Translation."
- Al Thani, D., Al Tamimi, A., Othman, A., Habib, A., Lahiri, A., & Ahmed, S. (2019, December). Mada Innovation Program: A Go-to-Market ecosystem for Arabic Accessibility Solutions. In 2019 7th International Conference on ICT & Accessibility (ICTA) (pp. 1-3). IEEE.
- Jemni, M., Semreen, S., Othman, A., Tmar, Z. and Aouiti, N., 2013, October. Toward the creation of an Arab Gloss for arabic Sign Language annotation. In Fourth International Conference on Information and Communication Technology and Accessibility (ICTA) (pp. 1-5). IEEE.



- Kipp M., Heloir A., Nguyen Q., Sign Language Avatars: Animation and Comprehensibility, International Workshop on Intelligent Virtual Agents, pp 113-126, 2011.
- Lagha, I. and Othman, A., 2019, December. Understanding Prosodic Pauses in Sign Language from Motion-Capture and Video-data. In 2019 7th International conference on ICT & Accessibility (ICTA) (pp. 1-4). IEEE.
- Othman, A. and Jemni, M., 2011. Statistical sign language machine translation: from English written text to American sign language gloss. arXiv preprint arXiv:1112.0168.
- Othman, Achraf, and Mohamed Jemni. 2012. "English-Asl Gloss Parallel Corpus 2012: Aslg-Pc12." In 5th Workshop on the Representation and Processing of Sign Languages: Interactions Between Corpus and Lexicon LREC.
- Othman, A. and Jemni, M., 2017, December. An XML-gloss annotation system for sign language processing. In 2017 6th International Conference on Information and Communication Technology and Accessibility (ICTA) (pp. 1-7). IEEE.
- Othman, A. and Jemni, M., 2019. Designing high accuracy statistical machine translation for sign language using parallel corpus: case study english and american sign language. Journal of Information Technology Research (JITR), 12(2), pp.134-158.
- Somers, H., 1999. Example-based machine translation. Machine translation, 14(2), pp.113-157.
- Tmar, Z., Othman, A. and Jemni, M., 2013, March. A rule-based approach for building an artificial English-ASL corpus. In 2013 International Conference on Electrical Engineering and Software Applications (pp. 1-4). IEEE.

