

BuHamad: The first Qatari virtual interpreter for Qatari Sign Language

Achraf Othman and Oussama El Ghoul
Mada Center

Abstract –

As well as all spoken languages, Sign Languages SL has structured grammar and syntax. Although it was visual, multi-dimensional and mainly based on gestures, SL follows specific grammatical rules. Consequently, the automatic generation of sign language should follow these rules. It is for this reason that Mada works on the development of a new framework that aims to support researchers and developers to create new innovative tools for the deaf. The objective is essentially to create tools that enhance the development of software using grammatically validated sentences. Signed languages are principally based on manual and facial gestures.

For this reason, the automatic generation of SL mainly deals with a 3D virtual signer. The use of avatars presents a mandatory task on SL generation. Since 2019, Mada has been working on designing and developing a personalized Qatari virtual sign language interpreter. The avatar is now published on the Mada website. It is used to translate the website content to Qatari SL. In this paper, we describe the project's objective, and we illustrate how it was designed.

Keywords: 3D signing avatar, Sign language generation, Qatari Sign Language

Introduction

Deaf communities in different countries employ different sign languages, which are not unified [1]. We found sign language (SL) in the United States, Australia, United Kingdom, France, Germany, Spain, Brazil, South Africa, China, Korea, Japan, and other countries. For Arab countries, we found that most of the existing research focused on the use of Arabic Sign Language or Unified Arabic Sign Language [2]. However, sign language differs from one Arab country to another, and Arab deaf communities use different sign languages, Such as Qatari Sign Language (QSL), Saudi Sign Language [3], Tunisia Sign Language [4]. To convey the information between deaf persons, the signer uses multiple channels to construct the Spatio-temporal environment [5], which makes the understanding and processing of sign language a challenge for researchers and opens as an active research field in different domains like linguistics [6], computational linguistics [7], and sociology [8]. From a computational linguistics angle, the computational sign language covers different disciplines:

- (1) sign language recognition [9],
- (2) construction and building of datasets in sign language,
- (3) machine translation from spoken language to sign language [10],

- (4) machine translation from sign language to sign language [11],
- (5) sign language synthesis using conversation agents or avatars [12],
- (6) sign language notation and annotation systems [13], and,
- (7) sign language generation [7] and production [14].

Sign language production, translation, and recognition require large-scale datasets and corpora to build accurate systems [15], [16]. However, there are no publicly available large-scale datasets in any sign language with high-quality annotation available for computational sign language processing. This limit impacts the progress of research and decreases the accessibility to information and technologies for the deaf communities and implies a false sense of technological readiness for them.

To address this issue, Mada Center launched the “Jumla Sign Language” research project that aims to improve the accessibility of deaf people by using 3D virtual signers. “Jumla Sign Language” project is composed of 5 elements:

- (1) the virtual sign language interpreter “Bu Hamad,” a signing avatar available and live on Mada website that translates web content to Qatari sign language in real-time, taking into consideration all components of sign language [17] (Figure 1)
- (2) the Mada Machine Translation to Sign Language that aims to provide a real-time translation of text in the Arabic language
- (3) the annotation sign language system [18]
- (4) an open-source repository containing scripts for body and facial recognition [18]
- (5) “Jumla Dataset” a library containing annotated videos.

In this paper, we focus on the first component (The virtual signer “BuHamad”).

The Avatar

This work aims to make a realistic Qatari Avatar that can automatically generate Qatari Sign Language. It is known that SL processing is an arduous task as it needs to be based on specific linguistics rules. However, many avatar-based translation systems use a word-by-word translation due to the lack of linguistic studies on Arabic SL [Ref]. Consequently, the lack of SDK and programming tools that help developers to build their applications.

According to many focus groups with Qatari deaf, the avatar should be realistic and produces natural movements and realistic facial expression. The designed avatar has a Qatari style wearing Qatari clothes Gutra and Thube. The creation of “BuHamad” requires solving many technical challenges. The avatar should support live and real-time animation. It should also be run on a web environment with low-speed internet connexion and low graphical resources. The skin is created by applying approximately 116.000 polygons and 71.000 vertices (Fig 2).



Figure 1: BuHamad avatar on Mada Website

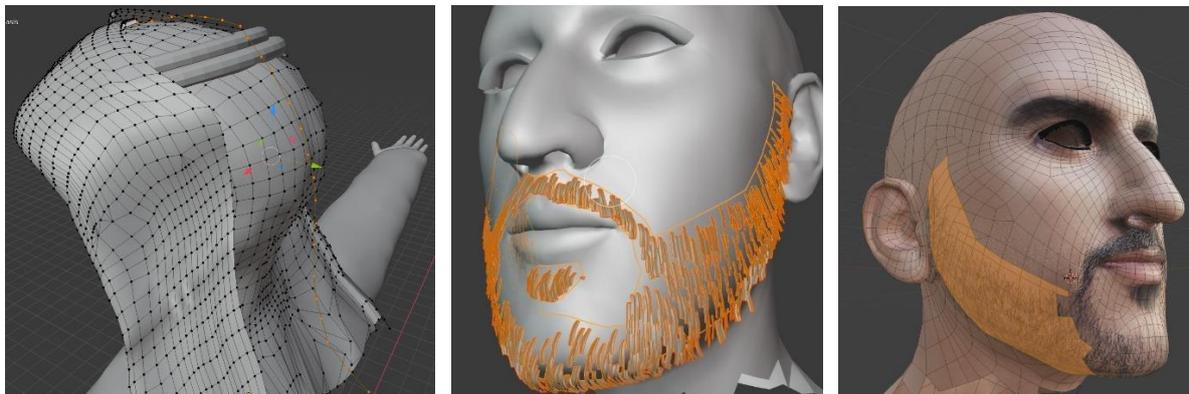


Figure 2: The design of the skin, beard, and clothes of the avatar

In order to animate the skin, we used a skeleton of 101 bones and applied 38 blend-shape to animate the face (Fig 3). The animation is just a description of changes along a timeline. For our 3D object, we adopted two ways of transforming its triangle mesh to create an animation:

- **Animation through skeletons** attached to the character: Rigid transforms are applied per bone. The skeleton was skinned to the skin of the character. By skinning, we mean the process of attaching the mesh of a 3D model to the virtual skeleton or rig so that when we move a character's rig, the skin moves with the rig, creating a quality character animation. This animation is used to make body and hand gestures.
- **Animation through the morphing of the mesh:** that is, moving each vertex of the mesh separately and storing its new location or by describing its change through particular functions. This technique is resources consuming. For this reason, we used it only to generate face animation.

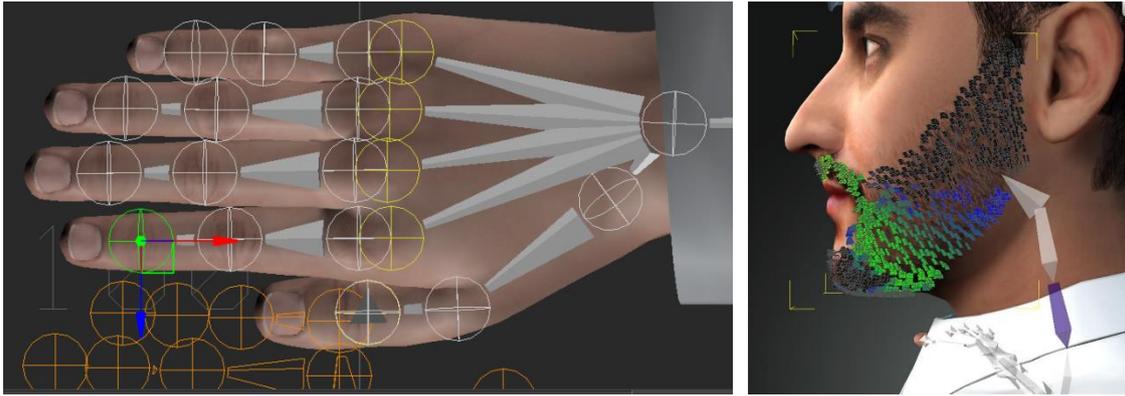


Figure 3: Skinning the model

We developed a set of tools and services to ensure the character's animation. This framework allows developers to easily integrate the character into their websites and translate texts to sign language.

Framework architecture

Mada *SLWebplayer* is designed to help developers to add signing avatars into their websites without the need to develop them from scratch. The Service is designed to guarantee high availability (using cloud servers) and extendibility of each component. We adopted 5 tier architecture organized as follows:

- *Website Server*: The web server contains the main website which includes the avatar.
- *Mada SLK Server*: The Server that provides the sign language toolkit proposed by Mada. The toolkit allows developers to integrate the avatar player into their websites, as described in the next section.
- *Mada RealTime SL Streamer*: Once moving the cursor over a paragraph, a dedicated script on SLK automatically generates a request to start receiving sign language animation as a compressed binary stream.
- *Signs Database Server*: A server containing the SL animations database (Words + Annotated Sentences)
- *Mada SLMT Server*: The Deep Learning Machine translation used to translate Arabic text to ArSL.

Network communication uses a protocol WebSocket. It provides a full-duplex communication channel over a single TCP connection. This protocol was standardized by the IETF as RFC 6455 in 2011, and the WebSocket API in Web IDL is being standardized by the W3C. The SL animation data is sampled and sent at regular time intervals, for which the length depends upon the configuration of the Mada Realtime SL Streamer. Typical sampling rates lie between 20 and 60 Hertz. The update rate of the real-time network stream can be modified separately.

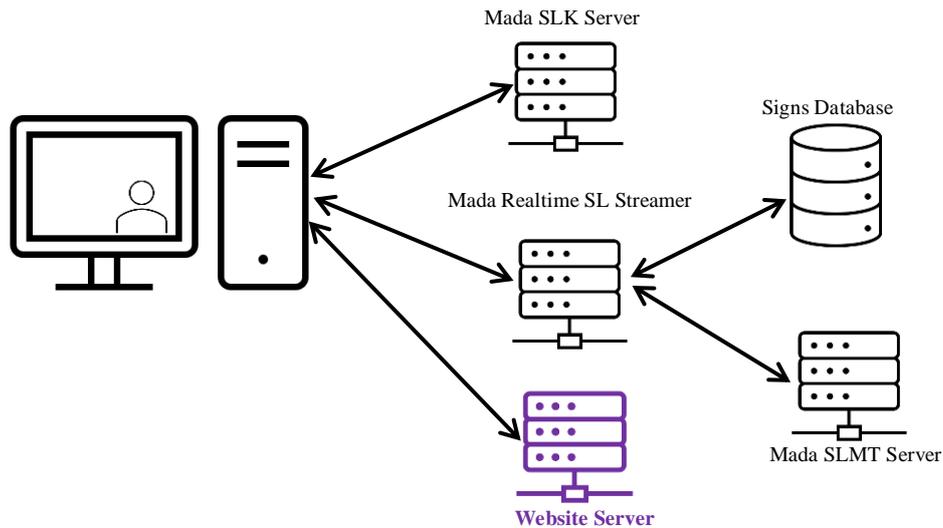


Figure 4: System architecture of the Virtual Qatari sign Language interpreter system

Conclusion

In this paper, we presented a new framework for integrating a 3D virtual signer into websites or desktop applications to make it accessible to the deaf by generating sign language. The proposed avatar is realistic, and it produces natural hand gestures and convincing facial expressions. The avatar is now online published on the website of Mada. The feedback of the deaf is looked promising. The work presents the first step of a big project that aims to improve the implementation of SL into websites.

Acknowledgment

We highly appreciate and are grateful for the team of the Mada Innovation Program and contributors from the Qatari deaf community, The Qatari Center of Social Cultural for the Deaf and Audio Education Complex. Mocap cleaning tasks and the Jumla sign language project have seen the light thanks to their support and contribution.

References

- [1] W. Sandler and D. Lillo-Martin, Sign language and linguistic universals. Cambridge University Press, 2006.
- [2] M. A. Abdel-Fattah, "Arabic sign language: a perspective," J. Deaf Stud. Deaf Educ., vol. 10, no. 2, pp. 212–221, 2005.
- [3] Y. O. M. Elhadj, Z. Zemirli, and K. Ayyadi, "Development of a bilingual parallel corpus of Arabic and Saudi Sign Language: Part I," in Intelligent Informatics, Springer, 2013, pp. 285–295.
- [4] M. Jemni, S. Semreen, A. Othman, Z. Tmar, and N. Aouiti, "Toward the creation of an Arab Gloss for arabic Sign Language annotation," in Fourth International Conference on Information and Communication Technology and Accessibility (ICTA), 2013, pp. 1–5.

- [5] T. Shanableh, K. Assaleh, and M. Al-Rousan, "Spatio-temporal feature-extraction techniques for isolated gesture recognition in Arabic sign language," *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, vol. 37, no. 3, pp. 641–650, 2007.
- [6] B. Woll, "The History of Sign Language Linguistics," *Oxf. Handb. Hist. Linguist.*, 2013.
- [7] D. Bragg et al., "Sign language recognition, generation, and translation: An interdisciplinary perspective," in *The 21st international ACM SIGACCESS conference on computers and accessibility*, 2019, pp. 16–31.
- [8] W. C. Stokoe, "Sociology in Sign Language Studies," *Sign Lang. Stud.*, vol. 36, no. 1, pp. 227–231, 1982.
- [9] R. Rastgoo, K. Kiani, and S. Escalera, "Sign language recognition: A deep survey," *Expert Syst. Appl.*, vol. 164, p. 113794, 2021.
- [10] J. Bungeroth and H. Ney, "Statistical sign language translation," in *Workshop on representation and processing of sign languages, LREC*, 2004, vol. 4, pp. 105–108.
- [11] A. Othman and M. Jemni, "Statistical sign language machine translation: from English written text to American sign language gloss," *ArXiv Prepr. ArXiv11120168*, 2011.
- [12] H. Maarif, R. Akmeliawati, and T. S. Gunawan, "Survey on language processing algorithm for sign language synthesizer," *Int. J. Robot. Mechatron.*, vol. 4, no. 2, pp. 39–48, 2018.
- [13] A. Schembri and O. Crasborn, "Issues in creating annotation standards for sign language description," in *sign-lang@ LREC 2010*, 2010, pp. 212–216.
- [14] R. Rastgoo, K. Kiani, S. Escalera, and M. Sabokrou, "Sign Language Production: A Review," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 3451–3461.
- [15] A. Duarte et al., "How2Sign: a large-scale multimodal dataset for continuous American sign language," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 2735–2744.
- [16] H. R. V. Joze and O. Koller, "Ms-asl: A large-scale data set and benchmark for understanding American sign language," *ArXiv Prepr. ArXiv181201053*, 2018.
- [17] "Mada – Digital Access for All." <https://mada.org.qa/> (accessed Nov. 30, 2021).
- [18] "Mada Center," GitHub. <https://github.com/madainnovation> (accessed Nov. 30, 2021).