

بو حمد: أول مترجم افتراضي قطري للغة الإشارة القطرية

أشرف عثمان وأسامة الغول
مركز مدى

ملخص

كما هو الحال مع جميع اللغات المنطوقة، فإن لغات الإشارة لها قواعد نحوية وبنية خاصة. على الرغم من أنها كانت مرئية ومتعددة الأبعاد وتستند أساسًا إلى الإيماءات، إلا أن لغة الإشارة تتبع قواعد نحوية محددة. وبالتالي، يجب أن يتبع الإنشاء الآلي للغة الإشارة هذه القواعد. ولهذا السبب، يعمل مركز مدى على تطوير إطار عمل جديد يهدف إلى دعم الباحثين والمطورين لإنشاء أدوات مبتكرة جديدة للضم. إن الهدف الأساسي هو إنشاء أدوات تعزز تطوير البرامج باستخدام جمل تم التحقق من صحتها نحويًا. وتعتمد لغات الإشارة بشكل أساسي على الإيماءات اليدوية وإيماءات الوجه.

ولهذا السبب، يتعامل التوليد الآلي للغة الإشارة بشكل أساسي مع شخصية افتراضية ثلاثية الأبعاد. حيث يمثل استخدام الشخصيات الافتراضية مهمة إلزامية في مجال توليد لغة الإشارة. ويعمل مركز مدى منذ عام 2019 على تصميم وتطوير مترجم افتراضي للغة الإشارة العربية والقطرية. وقد تم إطلاق هذه الشخصية الافتراضية الآن على الموقع الإلكتروني للمركز. حيث يتم استخدامها لترجمة محتوى الموقع إلى لغة الإشارة القطرية. وفي هذه الورقة، نصف هدف هذا المشروع ونوضح كيف تم تصميمه.

الكلمات المفتاحية: الشخصية الافتراضية ثلاثية الأبعاد للترجمة إلى لغة الإشارة، توليد لغة الإشارة، لغة الإشارة القطرية.

مقدمة

تستخدم مجتمعات الصم في البلدان المختلفة لغات إشارة مختلفة، وهي غير موحدة [1]. وقد وجدنا أن لغة الإشارة (SL) مستخدمة في الولايات المتحدة وأستراليا والمملكة المتحدة وفرنسا وألمانيا وإسبانيا والبرازيل وجنوب إفريقيا والصين وكوريا واليابان ودول أخرى. وبالنسبة للدول العربية، وجدنا أن معظم الأبحاث الحالية تركز على استخدام لغة الإشارة العربية أو لغة الإشارة العربية الموحدة [2]. ومع ذلك، تختلف لغة

الإشارة من دولة عربية إلى أخرى، حيث تستخدم مجتمعات الصم العربية لغات إشارة مختلفة، مثل لغة الإشارة القطرية (QSL) ولغة الإشارة السعودية [3] ولغة الإشارة التونسية [4]. وبهدف نقل المعلومات بين الصم، يستخدم مترجم لغة الإشارة قنوات متعددة لبناء البيئة المكانية الزمانية [5]، مما يجعل فهم لغة الإشارة ومعالجتها تحديًا للباحثين ويفتح مجال بحثٍ نشطٍ في مجالات مختلفة مثل اللغويات [6] واللغويات الحاسوبية [7] وعلم الاجتماع [8]. ومن جهة اللغويات الحاسوبية، فإن لغة الإشارة المحوسبة تغطي تخصصات مختلفة:

- (1) التعرف على لغة الإشارة [9] ،
- (2) بناء مجموعات البيانات بلغة الإشارة.
- (3) الترجمة الآلية من اللغة المنطوقة إلى لغة الإشارة [10]،
- (4) الترجمة الآلية من لغة الإشارة إلى لغة الإشارة [11]،
- (5) تركيب لغة الإشارة باستخدام وكلاء المحادثة أو الشخصيات الافتراضية [12] ،
- (6) أنظمة تدوين لغة الإشارة وإرفاق التعليقات التوضيحية [13]،
- (7) توليد وإنتاج لغة الإشارة [7] [14].

يتطلب إنتاج لغة الإشارة وترجمتها والتعرف عليها مجموعات بيانات واسعة النطاق ومجموعات لغوية لبناء أنظمة دقيقة [15]، [16]. ومع ذلك، لا توجد مجموعات بيانات واسعة النطاق متاحة للجمهور بأي لغة إشارة مع شرح توضيحي عالي الجودة لمعالجة لغة الإشارة المحوسبة. ويؤثر هذا الأمر على تقدم البحث ويقفل من إمكانية النفاذ إلى المعلومات والتكنولوجيا لمجتمعات الصم وينطوي على شعور زائف بالاستعداد التكنولوجي لهم.

ولمعالجة هذه المشكلة، أطلق مركز مدى مشروع بحث "لغة الإشارة جملة" الذي يهدف إلى تحسين إمكانية النفاذ للصم باستخدام شخصيات افتراضية ثلاثية الأبعاد للترجمة إلى هذه اللغة. ويتكون مشروع "لغة الإشارة جملة" من خمسة عناصر:

- (1) مترجم لغة الإشارة الافتراضية "بو حمد"، وهو متوفر للعمل مباشرة على موقع مركز مدى بحيث يترجم محتوى الويب إلى لغة الإشارة القطرية في الوقت الفعلي، مع الأخذ في الاعتبار جميع مكونات لغة الإشارة [17] (الشكل 1).
- (2) الترجمة الآلية إلى لغة الإشارة من مدى والتي تهدف إلى توفير ترجمة فورية للنص باللغة العربية
- (3) نظام إرفاق التعليقات التوضيحية للغة الإشارة [18].
- (4) مستودع مفتوح المصدر يحتوي على مكونات نصية للتعرف على الجسم والوجه [18].
- (5) "مجموعة بيانات جملة" وهي مكتبة تحتوي على مقاطع فيديو مرفقة بتعليقات توضيحية.

ونركز في هذا البحث على المكون الأول (مترجم لغة الإشارة الافتراضية "بو حمد").

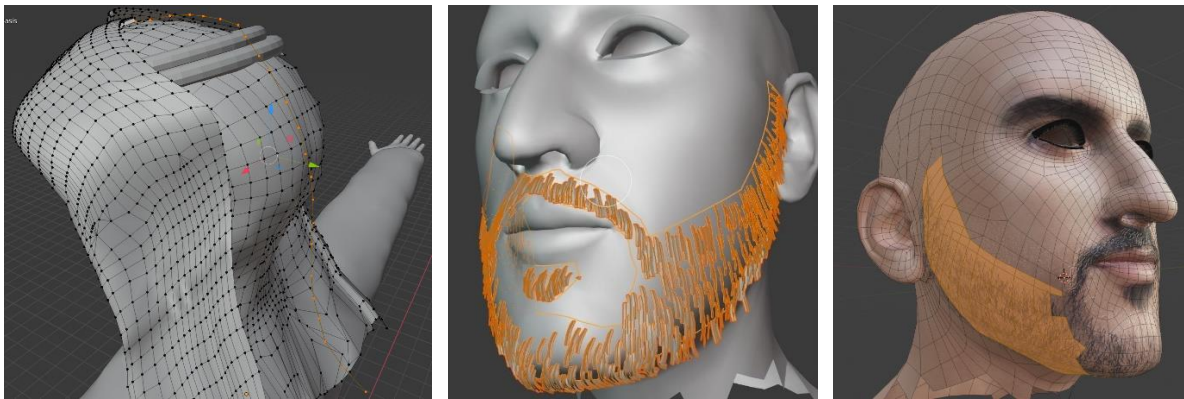
الشخصية الافتراضية ثلاثية الأبعاد

يهدف هذا العمل إلى إنشاء شخصية افتراضية قطرية واقعية يمكنها توليد لغة الإشارة القطرية بشكل آلي. ومن المعروف أن معالجة اللغة المصدر هي مهمة شاقة لأنها يجب أن تستند إلى قواعد لغوية محددة. ومع ذلك، فإن العديد من أنظمة الترجمة القائمة على الشخصيات الافتراضية تستخدم الترجمة كلمة بكلمة بسبب نقص الدراسات اللغوية على لغة الإشارة العربية [المرجع] وبالتالي، عدم وجود مجموعة أدوات تطوير البرمجيات وأدوات البرمجة التي تساعد المطورين على بناء تطبيقاتهم.

ووفقاً للعديد من مجموعات التركيز مع الصم القطريين، يجب أن تكون الشخصية الافتراضية واقعية تنتج حركات طبيعية وتعبيرات وجه معبرة وواقعية. إن الشخصية الافتراضية المصممة هنا تتمتع بمعالم شخص قطري يرتدي الملابس القطرية (الثوب والغترة). وقد تطلب تطوير شخصية "بو حمد" الافتراضية حل العديد من التحديات التكنولوجية، فهي يجب أن تدعم هذه الشخصية الرسوم المتحركة الحية والواقعية. ويجب أيضاً تشغيلها على بيئة ويب ذات اتصال إنترنت منخفض السرعة وموارد رسومية منخفضة. وفي الواقع فإن إنشاء جلد هذه الشخصية يتم من خلال بتطبيق ما يقرب من 116.000 شكل مضلع و 71.000 رأس زاوية (الشكل 2).



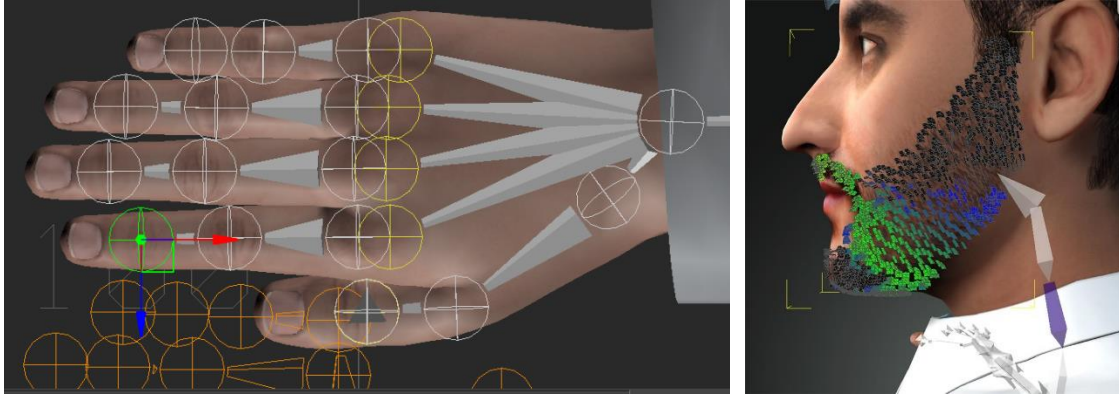
الشكل 1: المترجم الافتراضي " بو حمد " على موقع مدى



الشكل 2: تصميم بشرة ولحية وملابس الشخصية الافتراضية

ومن أجل تنشيط طبقة الجلد هذه، فقد استخدمنا هيكلًا عظميًا مكونًا من 101 عظمة وطبقنا 38 شكلًا مدمجًا لتحريك الوجه (الشكل 3). وتشكل الرسوم المتحركة مجرد وصف للتغيرات على طول الخط الزمني. وبالنسبة لهذه الشخصية ثلاثية الأبعاد، فقد اعتمدنا طريقتين لتحويل شبكة المثلث الخاصة به لإنشاء رسم متحرك:

- الرسوم المتحركة من خلال الهياكل العظمية المرتبطة بالشخصية: يتم تطبيق التحويلات الصلبة لكل عظمة. تم تصميم جلد الشخصية بما يتناسب مع الهيكل العظمي. ونعني بتطبيق الجلد على الهيكل عملية ربط شبكة نموذج ثلاثي الأبعاد بالهيكل العظمي الافتراضي بحيث أننا عندما نقوم بتحريك الهيكل العظمي لشخصية ما، يتحرك الجلد معه، مما يؤدي إلى إنشاء رسوم متحركة عالية الجودة للشخصية. وتستخدم هذه الرسوم المتحركة لعمل إيماءات الجسم واليد.
- الرسوم المتحركة من خلال تحويل الشبكة: أي تحريك كل رأس للشبكة بشكل منفصل وتخزين موقعها الجديد أو من خلال وصف تغييرها من خلال وظائف معينة. ولكن هذه التكنولوجيا تستهلك الموارد بشكل كبير. و لهذا السبب، فقد استخدمناها فقط لإنشاء رسوم متحركة للوجه.



الشكل 3: تطبيق الجلد على الشخصية

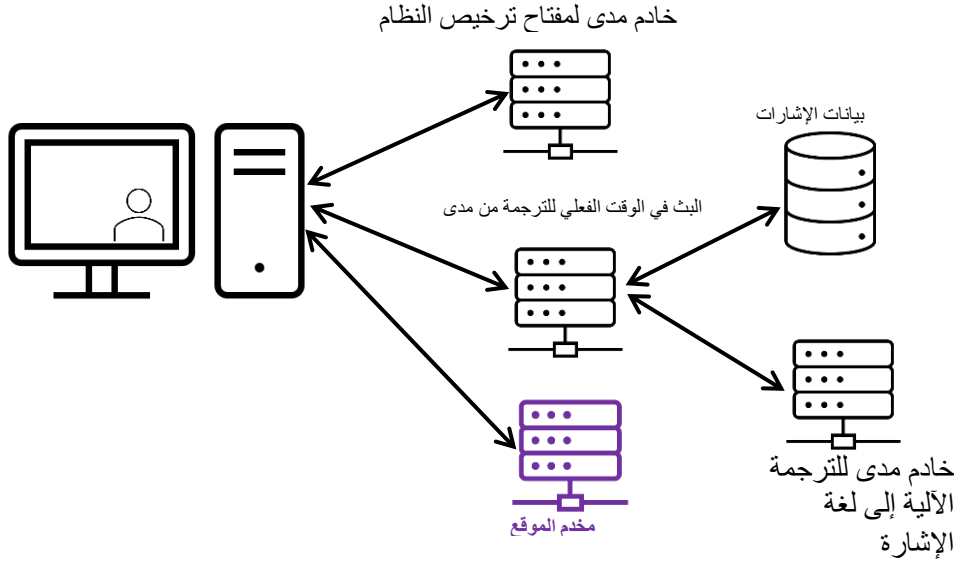
كما قمنا بتطوير مجموعة من الأدوات والخدمات لضمان نجاح تطبيق الرسوم المتحركة على الشخصية. ويتيح إطار العمل هذا للمطورين دمج هذه الشخصيات الافتراضية بسهولة في مواقعهم على الويب وترجمة النصوص إلى لغة الإشارة.

بنية الإطار

تم تصميم مشغل وسائط لغة الإشارة في مواقع الويب من مدى Mada SLWebplayer لمساعدة المطورين على إضافة مترجمين افتراضيين إلى لغة الإشارة إلى مواقعهم على الويب دون الحاجة إلى تطويرهم من الصفر. تم تصميم الخدمة لضمان توفرها الدائم (باستخدام الخوادم السحابية) وإمكانية تطوير كل مكون من مكوناتها. وقد اعتمدنا 5 طبقات معمارية منظمة على النحو التالي:

- خادم الموقع: يحتوي خادم الويب على موقع الويب الرئيسي الذي يتضمن الصورة الرمزية.
- خادم مدى لمفتاح ترخيص النظام SLK: الخادم الذي يوفر مجموعة أدوات لغة الإشارة التي اقترحها مدى. وهو يتيح مجموعة الأدوات للمطورين بما يسمح لهم بدمج مشغل الشخصية الافتراضية في مواقعهم على الويب، كما هو موضح في القسم التالي.
- نظام مدى لبث لغة الإشارة في الوقت الفعلي Mada RealTime SL Streamer: بمجرد تحريك المؤشر فوق فقرة ما، يقوم برنامج نصي مخصص على خادم مدى SLK تلقائياً بإنشاء طلب لبدء تلقي ترجمة لغة الإشارة عبر حركة الشخصية كمتوى بث ثنائي مضغوط.
- خادم قاعدة بيانات الإشارات: خادم يحتوي على قاعدة بيانات الرسوم المتحركة في لغة الإشارة (الكلمات + الجمل المشروحة)
- خادم مدى للترجمة الآلية إلى لغة الإشارة SLMT: الترجمة الآلية بميزة التعلم العميق المستخدمة لترجمة النص العربي إلى لغة الإشارة العربية.

تستخدم اتصالات الشبكة بروتوكول WebSocket والذي يوفر قناة اتصال ثنائية الاتجاه عبر اتصال TCP واحد. وقد تم توحيد هذا البروتوكول من قبل منظمة فريق عمل هندسة الإنترنت IETF كـ RFC 6455 في عام 2011، ويتم توحيد WebSocket API في Web IDL بواسطة رابطة الويب العالمية W3C. كما يتم أخذ عينات من بيانات الرسوم المتحركة للغة الإشارة وإرسالها على فترات زمنية منتظمة، حيث يعتمد طولها على تكوين نظام مدى لبث لغة الإشارة في الوقت الفعلي. وتتراوح معدلات أخذ العينات النموذجية بين 20 و 60 هرتز. كما يمكن تعديل معدل التحديث لدفق الشبكة في الوقت الفعلي بشكل منفصل.



الشكل 4: البنية المعمارية لنظام المترجم الافتراضي إلى لغة الإشارة القطرية

الخاتمة

لقد قدمنا في هذه الورقة إطاراً جديداً لدمج المترجم الافتراضي ثلاثي الأبعاد إلى لغة الإشارة في مواقع الويب أو تطبيقات سطح المكتب لجعله في متناول الصم عن طريق التوليد الآلي للغة الإشارة. إن الشخصية الافتراضية المقترحة هي شخصية واقعية تنتج إشارات يد طبيعية وتعبيرات وجه مقنعة. وقد قام مركز مدى بتفعيل هذه الشخصية الآن على موقعه الإلكتروني، حيث تبدو ردود الفعل من الأشخاص الصم واعدة حتى الآن. كما يمثل هذا العمل المرحلة الأولى من مشروع كبير يهدف إلى تحسين توفير الترجمة إلى لغة الإشارة في مواقع الويب.

شكر وتقدير

نحن نقدر تقديراً عالياً جهود فريق برنامج مدى للابتكار والمساهمين من مجتمع الصم القطري والمركز القطري الثقافي للصم و مجمع التربية السمعية، ونشعر بالامتنان لهم على مساهمتهم الفاعلة. حيث يعود نجاح مهام معالجة النقاط الحركة ومشروع "لغة الإشارة جملة" لدعمهم وتعاونهم.

المراجع

- [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).