Real Time Sign Language Translation Systems: A review study

Conference Paper · June 2022

DOI: 10.1109/MOCAST54814.2022.9837666

7 authors, including:

Maria Papatsimouli
University of Western Macedonia

Konstantinos Filippos Kollias
University of Western Macedonia

Lazaros Lazaridis
Region of Western Macedonia, Greece

George S. Maraslidis
University of Western Macedonia

Real Time Sign Language Translation Systems: A review study

Maria Papatsimouli, Konstantinos-Filippos Kollias, Lazaros Lazaridis, George Maraslidis,
Herakles Michailidis, Panagiotis Sarigiannidis and George F. Fragulis
Department of Electrical and Computer Engineering
University of Western Macedonia, Kozani, Greece
{*m.papatsimoulh, dece00063,dece00049,dece00079, dece00087,psarigiannidis,gfragulis*}@uowm.gr

***Abstract*—There are people who cannot communicate in the same way with others. Deaf and hard-of-hearing people use sign languages for their communication with other people. Sign languages are also used for the communication between deaf and non-deaf people, including different types of hand gestures and facial expressions for communication and emotional expression. Sign language recognition and gesture-based controls are applications that are used by gesture recognition technologies, and it is a fact that this technology has reduced the communication gap, while these systems are used for converting gestures to text or speech. The focus of our research is to analyze real-time sign language translators that are used for language translation. Sign Language Translation Systems that were developed from 2017 to 2021 are analysed in this paper.**

***Index Terms*—Sign Language, Sign Language Recognition, Handicapped aids, Application Program Interfaces, IoT.**

I. Introduction

Deaf and dumb people use sign language for communication, human interaction, and for understanding each other while communicating in a regular way [1].Nowadays, systems and software for sign language translation have been developed to translate text into animations enhancing the lives of deaf people, regarding communication and information access. During the last decade, there has been an increase in the usage of this technology. There is no difficulty when two deaf people communicate because they use sign language which is their common language. However, there can be difficulties when a deaf person wants to communicate with a non-deaf one [2].Hand movements recognition and translation, with all different approaches, is a topic that has plagued humanity for many years [3].

A. Languages

According to [4], *"a language is a complex system of communication with a vocabulary of conventional symbols and grammatical rules that are shared by members of a community and passed on from one generation to the next, that changes across time, and that is used to exchange an open-ended range of ideas, emotions and intentions."* Language is a communication system that offers arbitrary symbols such as traffic lights, monkey calls, or human language. Languages use words [5] and signs, such as symbols, and the link between form and meaning is signed, and spoken language may be arbitrary. The definition of arbitrary words or signs shows no link between their form and sense.

II. Sign languages

Deaf and dumb people have their own language, known as sign language, in which facial expressions, body and hand movements are used [6], [7]. Sign languages are recognized as natural human languages [8], [9] and are used as spoken languages, for social interaction, and communication with family [10]. However, there exist several differences between sign languages and spoken ones [11]. Sign languages are non-verbal languages utilising different hand gestures and facial expressions for communication, and emotional expression [12], [13]. Sign languages vary depending on the country and may even differ from area to area of the same region and have their own vocabularies and grammar [14], [15]. Finally, sign languages require many processes, such as hand configuration recognition, motion discrimination, identification of facial expressions, and recognition of linguistically relevant spatial contrasts [8], [5].

A. Applications of Sign language

One of the essential elements of sign language software applications is interaction. Interaction enables each user to be transformed from a passive recipient to an active member of learning process while keeping his interest undiminished. Sign Language is a visual language, and with the contribution of video, it can be included in any application in order to transfer information and provide hearing-impaired people with easy access to knowledge [16]. Various applications have been developed and are associated with the learning process and translation of sign languages from signs into spoken language or text.

B. Sign language translation systems approaches

Two main recognition mechanisms are used in sign language recognition systems i.e., sensor-based, and computer vision. Due to the technical image processing, a camera is used for capturing image or video. Hand movements are captured by a camera, and it is evident that a higher resolution camera needs more memory space and more calculation time and needs are required. Moreover, many high-performance techniques are required in computer vision technologies with more expensive sensors, making the application more expensive and complicated. The system has to be limited to a factual background without any noise or disturbance [17]. Besides, sign language recognition systems are categorized to: (i) data-gloves approaches and (ii) visual-based approaches.

C. Hand Gesture recognition categories

According to embedded sensors, there are two main categories of hand gesture systems: the camera-based and the wearable ones. The camera-based systems can reach high recognition efficiency and a high computational cost. Such systems are sensitive in some conditions, such as the background, lighting conditions, and the room's geometry, and constrained by the field of the camera's view. Also, on the sensor-based systems, accelerometers, gyroscopes, magnetometers and body sensor networks are used. Generally, wearable sensors are energy-constrained because micro-electromechanical technologies are used. In this system, data are collected from the onboard sensors and use either machine learning algorithms, mathematical models, fuzzy control techniques, or simple threshold-based algorithms to recognize signers' gestures. These systems have a relatively low price and are not sensitive to environmental conditions [18], [19], [20].

D. Data Gloves

A data glove is supplied with sensors and can directly acquire important data, such as finger bent degree, wrist orientation, and hand motion. The data glove is the input channel in these systems and transmits data to a mobile phone. There are some limitations regarding these systems. For instance, it is difficult for a user wearing a data glove to capture hand and finger movements, and the gloves can not acquire data such as the expressions of the face, lip-perusing, and the movements of the eyes. Moreover, a data glove can be affected by the environment, like user's location, background conditions, and collected data [21]. In the data-glove method, the signer wears an electronic glove with sensors which detect and transmit information. Most sign language translation systems on the market use the data-glove form because it is easy to acquire information about the degree of finger flexing of the hand. These types of systems require less computational power, and the translation is easier to be accomplished [22]. Furthermore, data-gloves can be expensive costing over $9000 US Dollars. It is also possible to use less expensive data gloves but they are usually more sensitive to noise and have a low number of sensors causing loss of important information and lower accuracy precision in translation. In addition to these, a data-glove is somehow less comfortable for the signers because the hands of each user have different sizes [22]. In table I, the advantages and the disadvantages of hand gloves are presented [22]. Data gloves can perceive fingerspelling and sign motions, which include static and dynamic signs. Generally, smart gloves involve wireless mode, which is more expensive, lightweight, reliable, and easy to use. This system consists of flex sensors, microcontrollers, and wireless transmitters [23]. The disadvantage of the data gloves is their cost, and a less expensive glove can be more vulnerable to noise. Also, a cheaper data glove may have a limited number of sensors, resulting in the loss of essential data and accuracy in the interpretation [23].

TABLE I

Advantages	Disadvantages
Wireless	Difficult to handle
Portable	Components are expensive

E. Visual based approach

Concerning the visual-based approach, there is a camera acquiring the images and videos that are utilised for the translation. The primary advantage of the visual-based methodology is frameworks' flexibility, including facial expressions, head movements, and lipreading. These systems are divided into two different approaches due to their operation, i.e., the hand-crafted shading gloves and the light of skin-color recognition. In the hand-crafted shading gloves procedure, the signer uses a color-coded glove [24], [23]. The colors give information from the image acquisition through color segmentation. These gloves are a regular pair of gloves with particular shading on every fingertip and palm and are inexpensive compared to smart gloves. The needed equipment is a webcam and color-shaded gloves, which are inexpensive. Webcams are used for image and video acquisition in RGB shading. The skin-color technique is the most common because signers communicate directly with the system, and the needed equipment is only a camera to capture the pictures. Further, a stop sign has to be done by the signer because the image acquisition continually runs [23]. The disadvantages of smart gloves have led to an increase on the application of the vision-based approach. In this approach, a camera is used to capture images of signers in order to recognize the sign language. The vision-based method has flexibility in the system, as these systems can be developed to include non-manual signals such as facial expressions recognition, head movements recognition, and perform lipreading. Additionally, there can be some problems including noise and complex algorithms for data processing [22].

F. Sign Language Translation Systems

Several projects have been developed for spoken language translation, the majority of which addressed translation in a limited amount of domains, such as tourism and medium-sized vocabularies [25]. Sign language translation systems are based on various types of statistical approaches [26], among other things, example-based methods [27], finite-state transducers [28], and other techniques. There is also 3D avatar animation that is an important technology in which virtual agents are embedded in spoken language systems and provide various services [25]. Gesture recognition in sign language translation has some limitations. The primary task is the collection of movement data. In sign languages, hands, full-body movements, and facial expressions are used for a perfect translation. Moreover, tasks such as data acquisition, processing, and recognition have to be simple for achieving a portable system.

These systems have to translate an extensive set of signs in real-time for supporting people with hearing loss daily, demanding powerful hardware that limits the portability of the systems. Signs can be static or dynamic. There are several algorithms for the translation and recognition of static signs, but there have been only a few attempts in the literature that dealt with dynamic gestures, attaining reasonable translation rates but within a particular test environment. Lastly, the grammar structure of sentences has to be reconstructed by sign language translation systems because sign languages appear to be different from their respective spoken ones [21].

G. Hand Gesture Recognition Process

Signers interact with computers by hand gesture recognition. Firstly, the input devices capture the hand region of the original images. Then, some features which can describe hand gestures are obtained. To retrieve the proper information, hand gestures are compared with the stored data in terms of similarity. When the region of interest is detected, features that are needed are retrieved. Finally, the output (text, voice, or video) is provided to the user [29]. This process is depicted in figure 1.

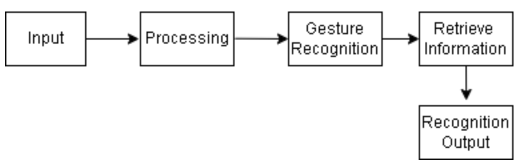

Fig. 1. Hand Gesture Recognition Process

III. Related Works in sign Language recognition using real-time approach

Various studies have researched Sign Language recognition using the real-time approach. There is a study [30] in which a project developed for educational purposes can translate static and dynamic signs of the American Sign Language. It is a game-learning application for students in learning applications, and the k-Nearest-Neighbour method was implemented as a classification method. An accuracy of 99.44% was achieved . According to [31], the "YSSA" is a wearable translator for the translation of the American sign language into English speech in real-time by using machine learning algorithms through a smartphone application. This system provides several features, including high sensitivity, low weight, high stretchability, and low cost. Moreover, it is based on contact electrification through a periodic contact area change of two different soft materials with opposite triboelectric polarizations, effectively converting a small tensile force or pressure into electricity. Additionally, the machine learning incorporation of the system achieved high recognition rates. However, the system's independence was limited as a mobile terminal is needed to listen to the speech. The [32] presented a glove that can translate the American Sign Language (ASL) alphabet into text using a computer or a smartphone. The packets with the needed information are sent serially to the user's PC to run a Python script. The glove uses strain sensors for comprising a piezoresistive composite of carbon particles embedded in a fluoroelastomer. Moreover, the sensors have a wearable electronic module consisting of digitizers, a microcontroller, and a Bluetooth radio. This glove translates all 26 letters of the ASL alphabet and costs less than 100 dollars. "Dastaana" [33] includes flex sensors furnished on the glove, with numerous hardware components centered on the user. Signers make signs according to American Sign Language, and after the processing, the data are transmitted to the application. According to the authors, many improvements can be made; for example, the interaction of gloves with devices across houses using IoT or SOS signals could be sent to the police stations or hospitals in case of emergency. Finally, [34] proposed an autonomous sign language recognition system that translates Indian Sign Language into speech. This system included various sensors like flex sensors, gyroscopes, and accelerometers for the determination of the position and orientation of the hand gesture.

TABLE II
Real Time Sign Language Translators

Paper	Name	Language	Year
[35]	–	B.S.L.	2021
[30]	LSTM-RN	A.S.L.	2021
[31]	YSSA	A.S.L.	2021
[33]	Dastaana	A.S.L.	2019
[32]	–	A.S.L.	2017
[34]	Talking Hands	I.S.L.	2017

IV. Conclusions

According to World Health Organization (WHO), more than 5% of the world's population has hearing loss, i.e., 466 million people, 432 million adults, and 34 million children. Deaf people also need a way to communicate and have developed a language that is directly accessible and useful for them that is called sign language [2]. Sign languages are the only languages that signers use for their communication in a familiar, effortless, useful, natural, and easy way. These problems prompted researchers to discover solutions for identifying sign-language communication and predict how two-way communication can be done. The deaf/hard-of-hearing people are also affected by these rapid changes. The use of technology reduces isolation, increases independence, and offers social, economic, educational, and other opportunities to deaf/hard-of-hearing people. The current survey presented six studies conducted from 2017 to 2021 concerning real-time sign language translation systems. Future research will provide a more comprehensive review of sign language translators depending on their embedded technology [38]. The prospective study of translators using IoT and machine learning technology needs also to be mentioned as it is better adapted to the user. In addition, sign language translators can be connected to various online CMS systems such as [36], [37], and [16] while enabling deaf people to take exams and gain skills.

V. Acknowledgment

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 957406.

References

[1] P. Dubey, “Sign language conversion flex sensor based on iot,” *International Journal of Research in Engineeringand Science (IJRES)*, vol. 9, no. 2, pp. 69–71, 2021.

[2] D. Sturman and D. Zeltzer, “A survey of glove-based input,” *IEEE Computer Graphics and Applications*, vol. 14, no. 1, pp. 30–39, 1994.

[3] Y. Wu and T. Huang, “Vision-based gesture recognition: A review. gesture-based communication in human-computer interaction,,” pp. 103–115, 1999.

[4] T. Johnston and A. Schembri, *Australian Sign Language (Auslan): an introduction to sign language linguistics*. United Kingdom: Cambridge University Press, Jan. 2007.

[5] K. Emmorey, *Language, cognition, and the brain: Insights from sign language research*. Lawrence Erlbaum Associates Publishers, 2002.

[6] P. S. Pooja Dubey, “Iot based sign language conversion,” *International Journal of Research in Engineering and Science (IJRES)*, vol. 9, pp. 84–89, 2021. [Online]. Available: www.ijres.org

[7] R. Sutton-Spence and B. Woll, *Linguistics and sign linguistics*. Cambridge University Press, 1999, p. 1–21.

[8] Z. Maalej, “Book review: Language, cognition, and the brain: Insights from sign language research,” *Linguist List*, vol. http://www.linguistlist.org/issues/13/13-1631.html, 01 2002.

[9] C. Monikowski, “Language, cognition, and the brain: Insights from sign language research,” *Studies in Second Language Acquisition*, vol. 26, no. 3, p. 497–498, 2004.

[10] W. Sandler and D. Lillo-Martin, *Sign language and linguistic universals*. Cambridge University Press, 2006.

[11] B. T. Tervoort, “Sign language: the study of deaf people and their language: J.g. kyle and b. woll, cambridge, cambridge university press, 1985. isbn 521 26075. ix+318 pp,” *Lingua*, vol. 70, no. 2, pp. 205–212, 1986. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0024384186900422

[12] P. Ambavane, R. Karjavkar, H. Pathare, S. Relekar, B. Alte, and N. K. Sharma, “A novel communication system for deaf and dumb people using gesture,” in *ITM Web of Conferences*, vol. 32. EDP Sciences, 2020, p. 02003.

[13] A. Das, L. Yadav, M. Singhal, R. Sachan, H. Goyal, K. Taparia, R. Gulati, A. Singh, and G. Trivedi, “Smart glove for sign language communications,” in *2016 International Conference on Accessibility to Digital World (ICADW)*, 2016, pp. 27–31.

[14] W. C. Stokoe Jr, “Sign language structure: An outline of the visual communication systems of the american deaf,” *Journal of deaf studies and deaf education*, vol. 10, no. 1, pp. 3–37, 2005.

[15] J. B. C. Christopoulos, “Sign language,” *Journal of Communication Disorders 1*, vol. 18, no. 1-20, 1985.

[16] M. Papatsimouli, L. Lazaridis, K.-F. Kollias, I. Skordas, and G. F. Fragulis, “Speak with signs: Active learning platform for greek sign language, english sign language, and their translation,” in *SHS Web of Conferences*, vol. 102. EDP Sciences, 2021, p. 01008.

[17] R. Wijayawickrama, T. P. Ravini Premachandra, and A. Chanaka, “Iot based sign language recognition system,” *Global Journal of Computer Science and Technology*, 2021.

[18] V. Sideridis, A. Zacharakis, G. Tzagkarakis, and M. Papadopouli, “Gesturekeeper: Gesture recognition for controlling devices in iot environments,” in *2019 27th European Signal Processing Conference (EUSIPCO)*, 2019, pp. 1–5.

[19] E. Garcia-Ceja, M. Z. Uddin, and J. Torresen, “Classification of recurrence plots’ distance matrices with a convolutional neural network for activity recognition,” *Procedia Computer Science*, vol. 130, pp. 157–163, 2018, the 9th International Conference on Ambient Systems, Networks and Technologies (ANT 2018) / The 8th International Conference on Sustainable Energy Information Technology (SEIT-2018) / Affiliated Workshops. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050918303752

[20] N. Twomey, T. Diethe, X. Fafoutis, A. Elsts, R. McConville, P. Flach, and I. Craddock, “A comprehensive study of activity recognition using accelerometers,” Mar. 2018.

[21] F. Pezzuoli, D. Tafaro, M. Pane, D. Corona, and M. L. Corradini, “Development of a new sign language translation system for people with autism spectrum disorder,” *Advances in Neurodevelopmental Disorders*, vol. 4, no. 4, pp. 439–446, 2020. [Online]. Available: https://doi.org/10.1007/s41252-020-00175-6

[22] M. C. Shubankar, B. and M. Priyaadharshini, “Iot device for disabled people,” *Procedia Computer Science*, vol. 165, pp. 189–195, 2019.

[23] A. Z. Shukor, M. F. Miskon, M. H. Jamaluddin, F. bin Ali@Ibrahim, M. F. Asyraf, and M. B. bin Bahar, “A new data glove approach for malaysian sign language detection,” *Procedia Computer Science*, vol. 76, pp. 60–67, 2015, 2015 IEEE International Symposium on Robotics and Intelligent Sensors (IEEE IRIS2015). [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050915037771

[24] R. Akmeliawati, M. P.-L. Ooi, and Y. C. Kuang, “Real-Time Malaysian Sign Language Translation using Colour Segmentation and Neural Network,” in *2007 IEEE Instrumentation & Measurement Technology Conference IMTC 2007*. Warsaw, Poland: IEEE, May 2007, pp. 1–6, iSSN: 1091-5281. [Online]. Available: http://ieeexplore.ieee.org/document/4258110/

[25] A. Aarssen, R. Genis, and E. van der Veeken, Eds., *A Bibliography of Sign Languages, 2008-2017 : With an Introduction by Myriam Vermeerbergen and Anna-Lena Nilsson*. Brill, 2018. [Online]. Available: http://library.oapen.org/handle/20.500.12657/37810

[26] I. G. Varea, F. J. Och, H. Ney, and F. Casacuberta, “Efficient integration of maximum entropy lexicon models within the training of statistical alignment models,” in *Machine Translation: From Research to Real Users*, S. D. Richardson, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 54–63.

[27] E. Sumita, Y. Akiba, T. Doi, A. Finch, K. Imamura, M. Paul, M. Shimohata, and T. Watanabe, “A corpus-centered approach to spoken language translation,” in *10th Conference of the European Chapter of the Association for Computational Linguistics*, 2003.

[28] F. Casacuberta and E. Vidal, “Machine Translation with Inferred Stochastic Finite-State Transducers,” *Computational Linguistics*, vol. 30, no. 2, pp. 205–225, 06 2004. [Online]. Available: https://doi.org/10.1162/089120104323093294

[29] S. Zhao, Z.-h. Chen, J.-T. Kim, J. Liang, J. Zhang, and Y.-B. Yuan, “Real-time hand gesture recognition using finger segmentation,” *The Scientific World Journal*, vol. 2014, p. 267872, 2014. [Online]. Available: https://doi.org/10.1155/2014/267872

[30] C. Lee, K. K. Ng, C.-H. Chen, H. Lau, S. Chung, and T. Tsoi, “American sign language recognition and training method with recurrent neural network,” *Expert Systems with Applications*, vol. 167, p. 114403, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417420310745

[31] Z. Zhou, K. Chen, X. Li, S. Zhang, Y. Wu, Y. Zhou, K. Meng, C. Sun, Q. He, W. Fan, E. Fan, Z. Lin, X. Tan, W. Deng, J. Yang, and J. Chen, “Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays,” *Nature Electronics*, vol. 3, no. 9, pp. 571–578, 2020. [Online]. Available: https://doi.org/10.1038/s41928-020-0428-6

[32] T. F. O’Connor, M. E. Fach, R. Miller, S. E. Root, P. P. Mercier, and D. J. Lipomi, “The language of glove: Wireless gesture decoder with low-power and stretchable hybrid electronics,” *PloS one*, vol. 12, no. 7, p. e0179766, 2017.

[33] S. B. Rizwan, M. S. Z. Khan, and M. Imran, “American sign language translation via smart wearable glove technology,” in *2019 International Symposium on Recent Advances in Electrical Engineering (RAEE)*, vol. 4, 2019, pp. 1–6.

[34] S. Y. Heera, M. K. Murthy, V. S. Sravanti, and S. Salvi, “Talking hands — an indian sign language to speech translating gloves,” in *2017 International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, 2017, pp. 746–751.

[35] T. Abedin, K. S. Prottoy, A. Moshruba, and S. B. Hakim, “Bangla sign language recognition using concatenated bdsl network,” *arXiv preprint arXiv:2107.11818*, 2021.

[36] G. F. Fragulis, M. Papatsimouli, L. Lazaridis, and I. A. Skordas, “An online dynamic examination system (odes) based on open source software tools,” *Software Impacts*, vol. 7, p. 100046, 2021.

[37] L. Lazaridis, M. Papatsimouli, and G. F. Fragulis, “A synchronous-asynchronous tele-education platform,” *International Journal of Smart Technology and Learning*, vol. 1, no. 2, pp. 122–139, 2019.

[38] G. Kokkonis, E. Gounopoulos, D. Tsiamitros, D. Stimoniaris, and G. F. Fragulis, “Designing interconnected haptic interfaces and actuators for teleoperations in mobile ad hoc networks,” *International Journal of Entertainment Technology and Management*, vol. 1, no. 1, pp. 43–63, 2020.