

Automated Bangla Sign Language Conversion System: Present and Future

Professor Dr. Mohammad Nurul Huda, Ishteaque Alam, Sabbir Arif Siddique

eGeneration Ltd.

mnh@cse.uju.ac.bd, ishteaque.ark@gmail.com, sabb.a.sidd@gmail.com

Abstract

Deaf and dumb people communicate with each other through sign language. In this research work we have analyzed and acknowledged useful features from most prominent sign language to Bangla text and speech language of classic and modern approaches.

Keywords: Bangla Sign Language, Bidirectional System, Static, Dynamic, Gesture Recognition, Real time Detection

Résumé

মূক ও বধির জনগোষ্ঠী পারস্পরিক যোগাযোগের জন্য ঈশারা ভাষা ব্যবহার করে থাকে। আমাদের এই গবেষণাটি ঈশারা ভাষা কে বাংলা কথিত ও লেখ্য ভাষায় রূপান্তর করার বর্তমান ও ভবিষ্যৎ প্রযুক্তি নিয়ে আলোকপাত করেছে।

1. INTRODUCTION

The method of human communication, either spoken or written, consists of the use of words in a structured and conventional way. But the deaf and dumb people cannot hear or speak. So, deaf and dumb people around the world communicate using sign language as distinct from spoken language in their everyday lives. A Sign Language is a visual language that uses a system of manual, facial and body movements as the means of communication. Sign language is not a universal language, and different sign languages are used in different countries. The Broca's and Wernicke's areas of left hemisphere of the brain process sign languages like other natural languages [1].

Sign language, any means of communication through bodily movements, especially of the hands and arms, used when spoken communication is impossible or not desirable. Wherever vocal communication is impossible, as between speakers of mutually unintelligible languages or when one or more would be communicators are deaf, sign language can be used to bridge the gap.

There are about 70 million deaf [2] hearing-impaired as well as hearing people in the world who use sign language as their first language or mother tongue. Each country has one or sometimes two or more sign languages, although different sign languages can share the same linguistic roots in the same way as spoken languages do. Though it is not an international language, the universal features in sign languages help to make it possible for users of different sign languages to understand one another far more quickly than users of unrelated spoken languages can. Deaf children must have rights for equal and quality education like all other children and expect that their educational rights are respected and supported by educational authorities.



Fig. 1. Signs for Bangla numeral

Hundreds of sign languages are being used around the world, which include, American Sign Language (ASL), Japanese Sign Language, (JSL), British Sign Language (BSL), Austrian Sign Language (GS), Bangladeshi Sign Language (BdSL) and so on.

Figure-1 shows the Bangla numeral sign [3] and Figure-2 shows the two handed Bangla vowels sign [4].



Fig. 2. Two handed sign images for Bangla vowels

Nuances within the deaf communities are the reasons for most of the differences in these signs of that area, and has led to a significant evolution of sign language worldwide.

All over the world, many systems had been developed by the researchers for different sign languages. An interactive sign language teaching program [5] was developed by Kadam et al. for American Sign Language (ASL) using glove. For detecting the hand gesture, they used flex sensors in the fingers of the glove. Verma et al. developed a system [6] for detecting sign language gesture using Microsoft Kinect. Use of Microsoft Kinect provides highly strong features form depth image which are very useful to make an accurate gesture recognition system. A bidirectional translation system [7] for Japanese Sign Language (JSL) designed by Imagawa et al. They made a real-time system for tracking hands of signer and find the location of each hand which is very necessary for detecting sign language gestures. An application named 'Gesture Audio Video Conferencing Application' proposed by Tarte et al. in

[8]. This application model includes translation of Naturally Spoken English Sentences to Visual Sign Language Gesture and vice versa. In this project input hand gesture of sign language can be detected by gesture template matching and it will produce text to speech output. On the other hand, natural voice can be detected by voice recognition system to fetch the input sentence and 61 producing output gestures for important words only.

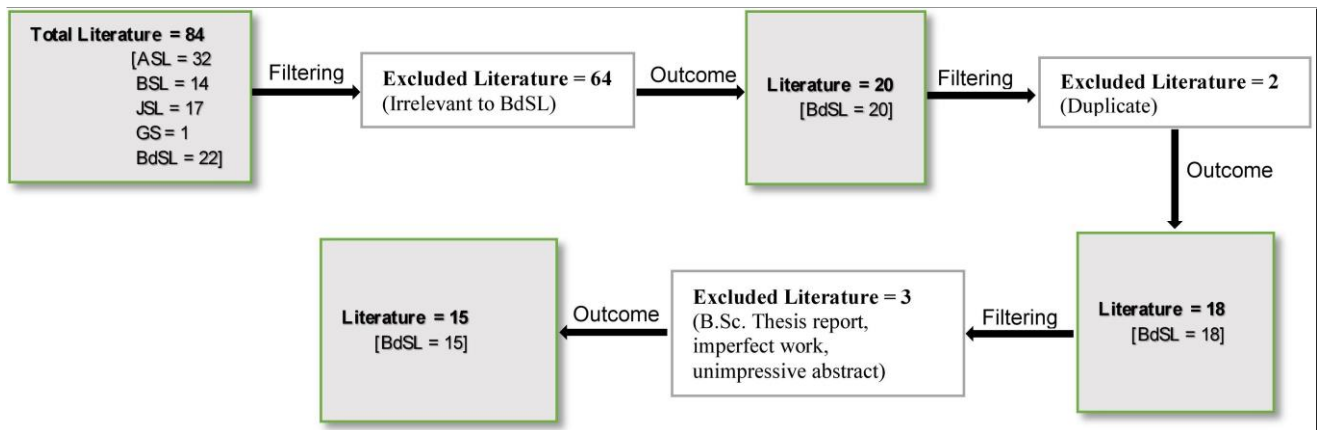


Fig. 3. Filtering articles to find all literature relevant to BdSL

Starner et al. developed two real-time hidden Markov model-based systems [9] for continuously detecting American Sign Language (ASL). Their first system used a camera on a desk which is actually second-person viewpoint for observing the signer's hand. The second system used a camera in a cap which was worn by the signer for observing his/her own hand gesture from first-person viewpoint. Using a 40 word lexicon, the first system achieved 92% accuracy while the second system achieved 98% accuracy.

The objective of the paper is to analysis the existing sign language translation system, a system that translate naturally Bangla spoken language to Bangla Sign Language and vice versa and then proposed a system that would have more flexibility and robustness for real life applications. For that purpose, we find out the previous related works which were developed for Bangla Sign Language in order to review and analyze these existing systems for developing a system to overcome the limitations of those.

Bangla is the most spoken language in Bangladesh and second in India. About 250 million native and about 300 million total speakers worldwide speaks in Bangla [10]. So a welldeveloped system for Bangla speech to Bangla Sign Language and vice versa is very much needed for communication with deaf people. While discussing further, we will use the term BL for naturally spoken Bangla Language and BdSL for Bangla Sign Language.

2. METHODOLOGY

To identify the best approach for a Sign Language translation system, one of the early steps is to review journal, articles, and conference papers related to this system. The best place for finding research papers are academic search engines, Google Scholar, Microsoft Academic Research, Directory of Open Access Journals, Science and Technology of Advanced Materials etc.

2.1 Search Methods

The results of search engines depend upon the selection of relevant keywords. In starting we used several different keywords, both alone and in combination. Later we narrow our focus and figured out which keywords describe our research topic best from the results of the searches. We also

used available filters in the academic search engines to filter down to our relevant research papers. The results of academic search engines come in the form of an abstract with a full citation (author, journal title, volume, page numbers, year, etc.). We read title and abstract from the results and also keep noticed to the journal or conference title to select research papers of our research topic of this study. Most frequently used keywords for searching journal, articles and conference papers for BdSL were "Sign Language", "Recognition of Sign Language", "Hand Gesture Recognition", "Sign Language Interpreter", "Real-time Gesture Recognition", "Communication System with Deaf and Dumb People", "Interpreter for Deaf and Dumb People", "Bidirectional Communication System for Sign Language". Later for finding more specific research paper about BdSL, we used the keywords like "Bangla", "Bengali", "Bangladeshi Deaf and Dumb People" as a prefix before the most frequently used keywords mentioned above.

2.2 Selection Criteria

We used some selection criteria to find the best fit articles related to this system. Articles of only reputed journals and conferences relevant to the topic were considered. The filtering process is represented by the Figure-3.

We have selected only the articles which were published within the year of 2000 to 2015. Out of all these research papers we targeted only those which are related to Bangla Sign Language. After all the filtering process, finally we have 15 research works related to BdSL.

The selection and filtering process can be summarized as:

- The literature must have discussed about Sign Language.
- Since the topic is about BdSL, the literature has to contain discussion or proposal or implementation about Bangla Sign Language.
- PhD or MS research works have been prioritized.
- Relevant, unique and informatics works are included
- Duplicate literature are excluded.

3. RESULT OF RELATED RESEARCH PAPERS IN BANGLA

In total we have collected 22 research papers related to sign language to speech conversion and vice versa. Among them 15 research papers are related to Bangla Sign Language. In this study, we tried to discuss about the methodologies used for Bangla Sign language to Bangla spoken language conversion and vice versa. The communication system with deaf and dumb people using sign language can be discussed as one way communication and bi-directional communication.

3.1 One Way Communication

One way communication system can be either conversion of naturally spoken Bangla Language (BL) to Bangla Sign Language (BdSL) or conversion of BdSL to BL. All those approaches are discussed below.

3.1.1 One Way Communication: Conversion of BL to BdSL

: A human machine interface named "Intelligent Assistant" designed by Eshaque et al. Which can understand only ten Bengali expressions. They used Microsoft Voice Command and Control Engine to capture sound input and convert it into recognizable specific text. Then the text is sent to match with previously stored words in the system's knowledge base. If the word is found in the database then the system displays a pre-stored 3D graphical hand gesture according to that text. The accuracy of "Intelligent Assistant" is 82% with only ten Bangla words. This system can understand only discrete words, not a whole sentence. They used one hand gesture but there are some signs in BdSL which requires both hands and facial expression too. Actually for understating only ten discrete words in Bangla, it doesn't require any intelligent system. Another system [12] developed by Sarkar et al. in 2009 which can translate an input Bangla sentence or word to the corresponding BdSL gestures using a set of rules. They took text input using either physical or on-screen keyboard. The translator dictionary is consist of 1000 words which were mostly taken from textbooks taught to children learning Bangla in the primary level. After taking an input and rearranging the words (if rearranging is necessary based on the rules), prestored video gestures of corresponding words are concatenated sequentially to produce a smooth output video in BdSL.

3.1.2 One Way Communication: Conversion of BdSL to BL

The conversion of BdSL to BL can be discussed based real time hand gesture recognition, one or two handed hand gesture recognition, static or dynamic gesture image recognition etc. Static gesture recognition means, a system can recognize sign gesture from a static image only. On the other hand dynamic gesture recognition means, the system can detect a sign gesture from real time video or any video input.

3.1.2.1 Real Time Dynamic BdSL Gesture Input

A geometrical model based hand gesture recognition system developed by Pavel et al. Which can analyze video

clips of different BdSL gestures, processing them and gives an audio output. This system calculates angles between different parts of the hand with body and then matches those values with pre-stored values in database which they created manually for different sign expressions. Then audio file of the corresponding gesture is played as output. Another work under the same authors in 2004 can detect and simulate 9 different gestures. They have achieved 97% accuracy by testing 9 gestures with 20 different students. "Intelligent Assistants for Speech Impaired People" developed by Rahman et al. which uses images of hand wearing glove as input. The glove contains different dots in each finger. The dots in the hand image are analyzed to detect which sign is shown by signer. Then pre-stored audio and written text output is shown corresponding to the sign. Only Bengali 1-10 numerals can be detected by this system. A real time computer vision based system made by Rahaman et al. that can detect two handed hand gesture showing sign from real time video recording and provides output for the corresponding BdSL sign. Only 6 Bengali vowels and 30 Bengali consonants can be recognized by this system. It was trained using 3600 images (36 different signs, 10 Signers-4 Female, 6 Male and 10 images for each). This system was tested using 3600 images of 10 different signers and achieved accuracy 98.17% for vowels and 94.75% for consonants. Another system developed by Jarman et al. in [3] can recognize 46 one hand gesture of Bengali alphabets and numerals. Multilayered feed-forward neural network with back-propagation algorithm was used to train the dataset. The system can perform rotation on BdSL gestures if it is required in some cases which is a good feature of this system. This system was tested using dynamic sign as 3 seconds videos where frames were extracted at a rate of 15 fps. The average recognition rate this system is 88.69%. As the system can perform sign detection from video clips which we called dynamic sign detection, it can easily be convert to perform in real time sign detection.

3.1.2.2 Recognition of Static Two Handed Hand Gesture

An intelligent sign language verification system offered by Rahim et al. used image processing, clustering and neural network concepts. The database of this system contains clustering information for the detection of gestures. Useful features are generated from input image by image processing technique. These features are classified by Neural Networking process according to the clusters information. Then the output is shown both in visual and textual form. This system was tested on BdSL but it can any sign language with prior image processing and clustering. Not only two handed hand gesture, it is even capable of deal with other body parts according to the training. This system is called "Intelligent System" because it can train itself automatically. They achieved 92% (approx.) accuracy by using 24 different sign images with 10 clusters. Using artificial neural network, Rahman et al. designed a system described in. This system does not require any gloves or visual marking system, it just requires image of bare hand. It can recognize 36 letters of BdSL alphabet with an accuracy of 80.902%. Deb et al. developed a system [4] which used two different color wristbands to easily remove the forearm in binary image. After detecting the hand sign from the candidate region, this system used template matching with the use of normalizes cross-correlation to recognize different types of hand sign. This

system can detect Bangla Sign Language gestures of 10 Bangla alphabets with 96% success rate. For detecting two handed hand gesture, Yasir et al. designed a system which can recognize Bangla sign images of 15 Bangla letters. They used machine learning techniques PCA and LDA, neural network for training and testing purpose of their system. To find out a better solution they used three different models. In first model, input images were trained and tested using neural network only and resulted output comes with an accuracy of 67.15%. In the second model, Principal component analysis was used to train the dataset and resulted output had an accuracy of only 26%. And in last model LDA algorithm was used for feature extraction and resulted output achieved an accuracy of 100%.

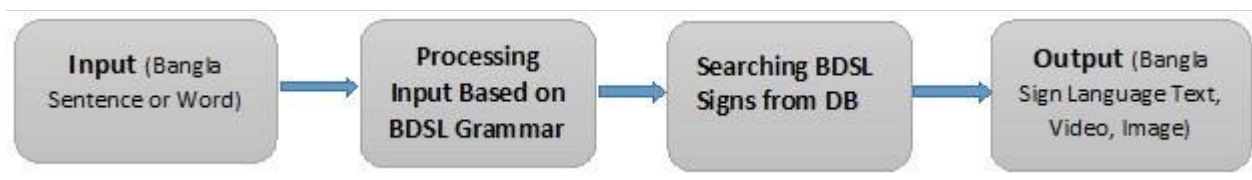
3.1.2.3 Recognition of Static Single Hand Gesture:

A fuzzy rule based system designed by Ayshee et al. to recognize static single hand gesture for only two Bengali letters. Angles of fingers are calculated to determine the hand configurations which are basically the fuzzy rules to match them with predefined fuzzy rules for two Bengali letters. In this system, static image of single hand gesture can be taken as input from one viewpoint only. All the fingers may not be visible from one viewpoint which is a limitation of this system. Karmokar et al. developed a system in by using Neural Network Ensemble for the recognition of single hand gesture of BdSL. This system was trained for 47 signs of BdSL including Bengali alphabets and numerals. They achieved 93% accuracy using negative correlation learning with 10 neural networks with feature extraction.

3.1.2.4 Recognition of Static Image of BdSL:

For recognition of static images, a system in developed by Rahim et al. which can take input as static sign gesture image and find out grammatical similarity with previously learned images. This system used back propagation algorithm of artificial neural network for learning and detection purpose. With 2000 learned images, the system was tested and the output achieved 99% accuracy.

3.2 Bi-directional Communication:



Language and (b) Bangla Sign Language to Bangla Language.

4.1 Bangla Language to Bangla Sign Language

The basic steps for the conversion of Bangla Language to Bangla Sign Language can be divided into four parts. Those are taking input as text, processing input based on predefined grammatical rules, searching processed text into database, preparing output and display it in monitor. The whole process can be simply represented by the Figure-4. First of all we have to construct a database containing all Bangla alphabets, numerals, words used in BdSL and mapping them with their corresponding BdSL gestures. BdSL gestures can be made as a mini video clips for each sign.

Sentence making grammar for Bangla Language and Bangla Sign Language has a huge difference. Both of them follow their own grammatical rules for making a complete understandable sentence. So for the conversion of Bangla Language to BdSL and vice versa, we have to define some rules for this type of conversion and implement it properly. For learning rules and implement it, we can use machine learning process. Rules based on Bangla Parts of Speech can also be implemented.

The input of Bangla Language can be either voice input using microphone or text input using physical or onscreen keyboard. Still there is no well-developed Bangla voice recognition system available and to stay focused in our main purpose, we will simply use textual input of Bangla language.

The output will consist of two parts. One part will show a textual output of grammatically converted BdSL from Bangla Language if the conversion is needed. Other part will show the corresponding BdSL gestures output of that converted BdSL text. The BdSL gestures will be represented by sequentially joining pre-stored video clips of corresponding words in that BdSL text.

Fig. 4. Conversion of Bangla Spoken Language to BdSL

A bidirectional system proposed by Datta et al. based on their previous work. This two way communication system includes Bangla Text to BdSL and BdSL to Bangla Text conversion. They did not do any implementation work for this proposed two way communication system but their previous work has an implementation for the conversion of Bangla text to BdSL gestures.

4. PROPOSED SYSTEM

For fluent communication with deaf people who use BdSL, our proposed system contains two major parts. This system includes conversion of (a) Bangla Language to Bangla Sign

So the overall process is, we will take Bangla Language as textual input. The input text will rearranged by the grammatical rules of Bangla Language to Bangla Sign Language conversion if the conversion is needed to perform. The words in the converted BdSL text will be searched in the database to extract their corresponding gesture video clips. After sequentially joining the video clips for each words in that text, both the textual and video output for BL to BdSL will be shown in the output screen. A similar process has been developed in [12] which has a translator dictionary containing 1000 words. A complete set of words used in BdSL need to add with their

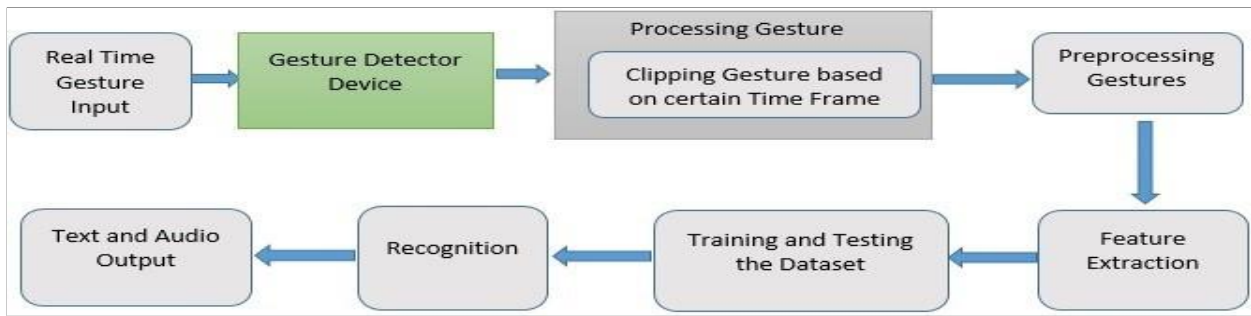


Fig. 5. Conversion of BdSL to Bangla Spoken Language

corresponding BdSL gestures. They implemented few rules for the conversion of BL to BdSL where many more rules need to be added.

4.2 Bangla Sign Language to Bangla Language

The target of BdSL to Bangla Language conversion in our proposed system is to detecting real time continuous dynamic signs using gesture detecting devices like leap motion or Kinect sensor and producing output in Bangla Language. Major parts of this conversion are taking real time gesture input, processing the gesture to extract features, training and testing input data, finally producing the recognized text and audio for the input Bangla gesture sign. Conversion of Bangla Sign Language to Bangla Language can be represented by the Figure-5.

The main purpose of this proposed model is to detecting consecutive Bangla signs in real time. The input real time gesture can be clipped based on a fixed time frame. If we use certain time frame to clip the gesture input then we have to search for a single unique sign in that sign gesture. The gesture sign will pass through the preprocessing step. Only hand signs are considering for the detection of BdSL, so we have to detect the hand region from extracted gestures, crop that hand region. Then important features have to extract from that gesture sign. Initially we have to train our dataset using those features to detect various BdSL gestures. After the training session, an input hand gesture sign can be recognized by comparing its features with the trained dataset. The output will be shown as both in textual and audio format for the corresponding input Bangla gesture sign.

5. CONCLUSION

This paper presents a comparison among the research works had done for development of the communication system between Bangla Sign Language and Bangla Language. It also proposed a two way communication system for BdSL to Bangla Language and vice versa. We focused on real time consecutive dynamic Bangla sign recognition for BdSL to Bangla Language conversion. On the other hand, Bangla Language to BdSL conversion needs a huge dataset of Bangla words mapped with their corresponding Bangla signs. A large set of grammatical rules has to develop and implement for proper conversion of Bangla Language to BdSL. Such a welldeveloped two way communication system is not still available for the deaf people who use BdSL in their regular life. In this system, we tried to pick the best solutions for the

limitations identified in the previous works and combine them with our own proposal.

Reference

- [1] R. Campbell, M. MacSweeney, and D. Waters, "Sign language and the brain: a review," *Journal of Deaf Studies and Deaf Education*, 2007.
- [2] "Sign language - wfd — world federation of the deaf," <http://wfdeaf.org/human-rights/crpd/sign-language>, (Visited on 09/30/2015).
- [3] N. A. M. J. I. Angur M. Jarman, Samiul Arshad, "An automated bengali sign language recognition system based on fingertip finder algorithm," in *International Journal of Electronics & Informatics (IJEI)*, ser. 1, vol. 4. Center for Natural Science & Engineering Research, 2015.
- [4] K. Deb, H. P. Mony, and S. Chowdhury, "Two-handed sign language recognition for bangla character using normalized cross correlation," *Global Journals Inc.(USA)*, vol. 12, 2012.
- [5] K. Kadam, R. Ganu, A. Bhosekar, and S. Joshi, "American sign language interpreter," in *Technology for Education (T4E)*, 2012 IEEE Fourth International Conference on. IEEE, 2012, pp. 157–159.
- [6] H. V. Verma, E. Aggarwal, and S. Chandra, "Gesture recognition using kinect for sign language translation," in *Image Information Processing (ICIIP)*, 2013 IEEE Second International Conference on. IEEE, 2013, pp. 96–100.
- [7] K. Imagawa, S. Lu, and S. Igi, "Color-based hands tracking system for sign language recognition," in *Automatic Face and Gesture Recognition*, 1998. Proceedings. Third IEEE International Conference on. IEEE, 1998, pp. 462–467.
- [8] N. T. Tarte, S. N. Bhadane, and P. S. Kulkarni, "A gesture audio-video conferencing application for the ease of communication between normal person and deaf and dumb person," *Int. J. Sci. Eng. Technol. Res.*, vol. 3, no. 5, p. 14871490, 2014.
- [9] T. Starner, J. Weaver, and A. Pentland, "Real-time american sign language recognition using desk and wearable computer based video," *Pattern Analysis and Machine Intelligence*, *IEEE Transactions on*, vol. 20, no. 12, pp. 1371–1375, 1998.
- [10] "Bengali language - wikipedia, the free encyclopedia," https://en.wikipedia.org/wiki/Bengali_language, (Visited on 09/30/2015).