# Motion synthesis for automatic animation of sign language

Jong Ho Jeong[1], Hee Jae Hwang[2], Hong Nyeom Sung[3] and Chil Woo Lee[4]

[1, 2, 3] Department of Computer Engineering, Chonnam National University, Korea
[4] Department of Software Engineering, Chonnam National University, Korea
[1]wkrlsk23@gmail.com, [2]certificate@kakao.com, [3]6590f1@naver.com, [4]leecw@jnu.ac.kr

**Abstract.** In this paper, we describe an algorithm and system that automatically animates a given Korean sentence into sign language. We plan to use this technology directly in cultural facilities such as museums, exhibition, and performance halls to improve the quality of life of the hearing-impaired. To achieve this goal, we made a sign language dictionary composed of about 2,500 words by capturing the sign language motion of the hearing-impaired and develop a technology to automatically generate sign language animation based on Unity. However, since all situations cannot be directly expressed only with the words recorded in the dictionary, it is necessary to create a sign language action of new meaning by combining two or more actions. In this paper, we describe the overview of the system developed so far and the connection of continuous sign language motions for expressing various sign language sentences and the synthesis algorithm of both arm motions in detail.

**Keywords:** Sign Language, Unity, automatic Animation algorithm, motion synthesis

## 1 Introduction

As information and communication technology develops, we use many useful services in our daily lives. However, the disabled people who have unwell body or difficulties in seeing and hearing often do not benefit from such high-tech services. In particular, the inconvenience that the hearing-impaired people have in using cultural facilities such as museums, exhibition halls, and performance halls is very serious. [1] In this study, to solve the difficulties of hearing-impaired people's experience when using cultural facilities, we describe a sign language animation automation system that translates Korean sentences into sign language sentences and then converts sign language sentences into sign language.

Since sign language has a different grammatical system from Korean, [2] it is treated as a separate language, just like foreign languages. Therefore, an independent Korean sign language dictionary has been defined as a separate language. But it's currently urgent to supplement and standardize the dictionary because not only the number of words is small, but also the expression behavior of one sign language word varies depending on the region or individual.

To automatically generate Korean sign language in cultural facilities, it is necessary to solve the following problems. The first is the definition and standardization of terms used in exhibition and performance halls. The terms used in exhibition halls are highly specialized, and there are many things that have not yet been defined in standard sign language. Therefore, definition and standardization of difficult terminology must be preceded. Second, since Korean sign language has a completely different gramma system from Korean, it is necessary to develop 'Korean'-'Korean sign language' automatic translation technology. As artificial intelligence technology develops, technology for automatically translating foreign languages into Korean or Korean into foreign languages has been developed and is widely used in everyday life, but the technology for effectively translate Korean to Korean sign language is still insufficient. Thirdly, technology is needed to convert translated Korean sign language sentences into Korean sign language animations. In other words, with an input data composed by Korean sentence, a technique for playing the sign language with a continuous animation suitable for the Korean sign language gramma system is required. We are trying to solve these three problems through joint research “development of intelligent exhibition commentary text/Korean sign language conversion technology for the hearing impaired” supported by the Ministry of Culture, Sports, and Tourism of Korea. In this paper, we describe the natural connection method of continuous motions which is the core of an algorithm, that converts a given sign language sentence into sign language animation, mentioned in third problem.

If one sentence is composed of several sign language words and the sign language actions corresponding to the words are accurately defined in the sign language dictionary, one sign language sentence can be completed by continuously connecting the words in the sign language dictionary. However, since the motion data for each word defined in the dictionary always includes the start and end motions for each word to express it as an independent word, simply connecting the words in sentence does not naturally create a continuous motion, so it cannot express the original meaning. In other words, connecting several independent motions into one natural motion without discontinuity becomes the most important factor in automatic sign language generation.

Sign language words have different meanings depending on the shape of the hand and fingers, the direction of the palm, the position of the hand, and the movement of the hand [3]. Therefore, if the left and right Sign language express independent words with different meanings, various Sign language can be expressed by the combination of the two Sign language. Since we cannot record all the Sign words using in our daily life in a sign language dictionary; That is, since we cannot record all Sign motions as motion capture data, it is necessary to combine independent motions and use them as new words or other expressions. In this paper, we describe an algorithm that continuously connects independent words recorded in the Korean sign language dictionary, synthesizes two words with different meanings to create a new compound word, and automatically creates a sign language animation using it. Through this algorithm, the number of sign language words can be increased, various and natural sign language animations can be displayed.

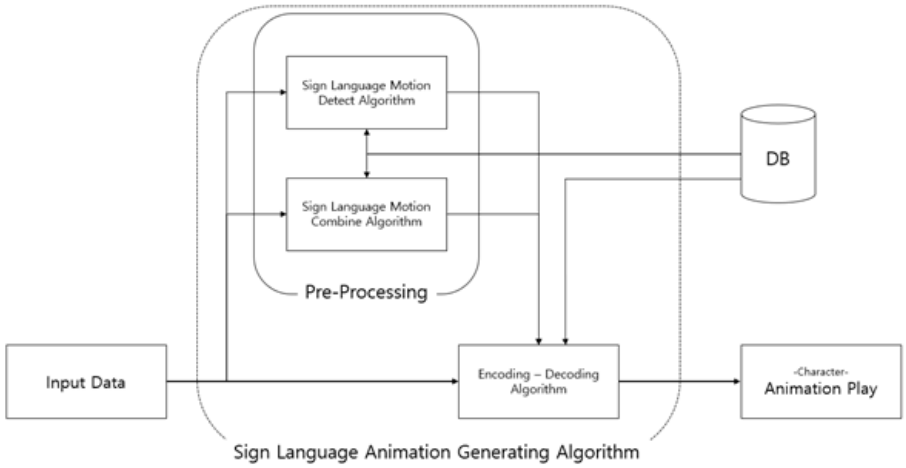


**Fig. 1.** Korean Sign Language automatic generation system Algorithm diagram. Before encoding-decoding algorithm performed, pre-processing is performed if necessary.

## 2 Sign language automatic generation system

### 2.1 Overall system overview

Sign language automatic generation system goes through the process shown in Figure 1. First, system analyze the input sentence to determine whether preprocessing is needed or not. If it is determined that pre-processing is necessary, a pre-processing step is performed. In the pre-processing step, system obtain the information which is necessary in final animation step by playing the sign language motion, synthesize the sign language motion which is necessary in final animation step. By synthesizing sign language motion in advance and using the motion data as it is in the final animation process, natural motion synthesis is possible. Therefore, the pre-processing process is divided into a sign language motion detection process and a sign word synthesis process.

After the pre-processing step, the animation step is executed. In the animation step, the Korean sign language sentence input using the pre-captured data is reproduced through and encoding-decoding algorithm. In this step, algorithm reproduce facial expression and sign language motion at an appropriate position input by user to make a natural sign language expression based on input sentence.

# 3 Representation of sign language word data

## 3.1 Korean sign language dictionary

To automatically generate sign language, it is first necessary to translate Korean sentences into sign language. In this study, sign language translation is performed by deep learning [4], and the result of translate is used as input to the system. As shown in Figure 2, in the entire process, we use the sign language dictionary produced through this study. This dictionary is divided into several categories to suit the characteristics of the word. In detail, There is SUJI-words (sign language word) that can be expressed with hand gestures, BISUJI-words (non-manual signal) that cannot be expressed by hand, such as facial expressions, JIWHA-words (finger language word) that written directly with finger, such as numbers, and SUHYUNG-words (Hand Style word) that used to synthesize two sign-language words. Each word has a category number to which the word belongs, a unique number of the word, and motion data in FBX format. Simply speaking, this dictionary can be said to be a database that systematically links sign language motion data to Korean words. This dictionary is used to find out if input data is a new word, or to generate an animation of input sentence. In this study, about 2500 words were divided into categories, and the sign language motion performed by the hearing-impaired for each word were captured as digital data to create a sign language dictionary. Since sign language is expressed as one independent motion that expresses meaning unlike general language, one sentence can be completed by continuously connecting motions corresponding to each word.

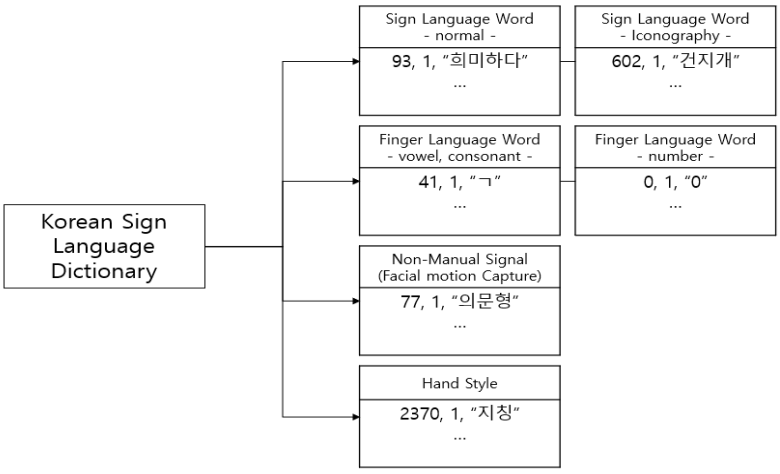


**Fig. 2.** Korean Sign Language Dictionary composition. The Dictionary compose 4 categories. In case of Sign language, sign language used in special situation such as museum are classified as iconography, and the others are classified as general. In case of Finger language, it is distinguished by finger language refers to vowels, consonants of Korean, and finger language refers to numbers.

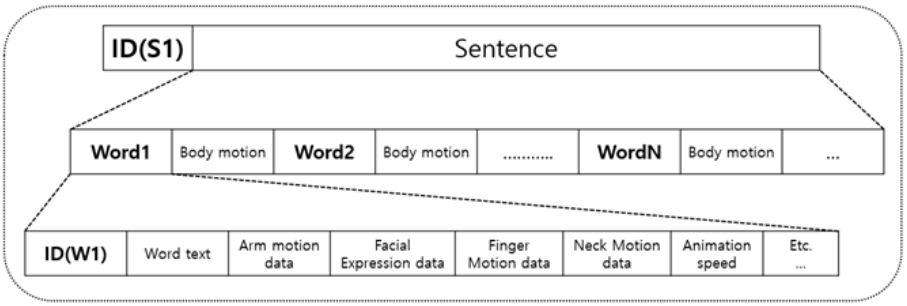


**Fig. 3.** Sign Language Sentence and Word Description Model. One Sentence contains several words, and one words contains various variables.

### 3.2 Korean sign language descriptor

The automatic sign language generation system reproduces Korean sentences translated into sign language words as sign language animations [5][6]. At this time, the sign language descriptor shown in Figure 3 is used to accurately reproduce the input sign language sentence as an animation. According to the sign language descriptor, one sentence data consists of several words, and various variables for animation control are included inside one word. In the system proposed in this paper, when playing an animation using the encoding-decoding algorithm in the animation playback stage, natural animation can be created at the correct timing by using the variables that words have.

## 4 An Algorithm for Automatic Sign language generation

### 4.1 Pre-processing for motion combination

**Sign language animation automatic generation algorithm**

*Pre-processing for motion combining.*

The sign language animation generation algorithm is preprocessed through two processes. First, in the sign language motion detection process, information necessary for connecting two motions is inspected to create one animation data, and in the second step, two words are synthesized to create a new word.

#### 4.1.1 Connection of two consecutive sign language motions

Sign language words store motion representing the meaning of each word as digital data. When acquiring motion data using a motion capture device, motions at the start and end of words are inserted to distinguish motions. Therefore, to create a natural sign language animation, it is necessary to remove the motion used for word classification [7].

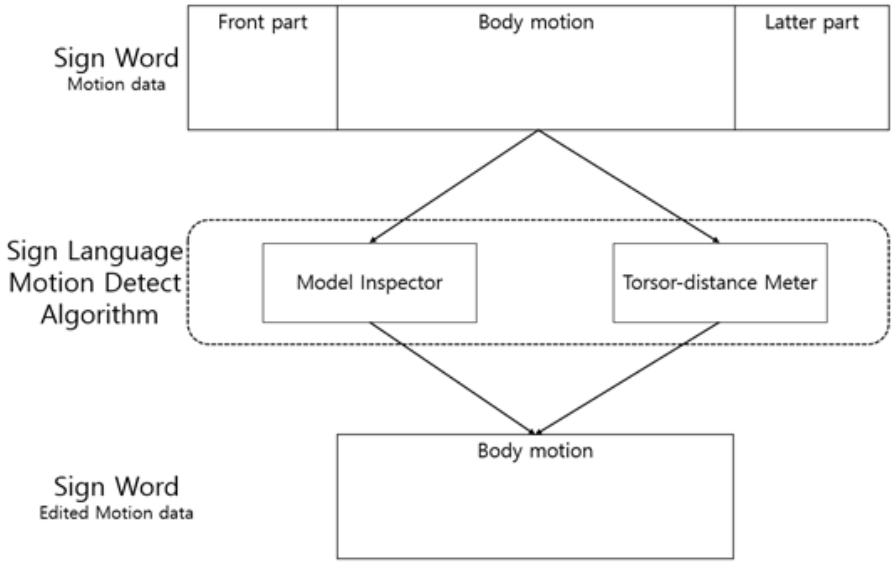


**Fig. 4.** Sign Language Motion Detect Algorithm diagram. 2 modules work in Sign Language Motion Detect Algorithm to inspect motion used for motion classification. After inspect Motion data, Sign language word edited.

The algorithm in Figure 4 is used to cut the back-and-forth motion of the motion data. In this algorithm, the Model Inspector, which reproduces sign language motions and detects the position of the hand and the torsor-distance Meter module, which calculates the distance to the torsor, operate. Model Inspector designates a specific location and records the current time when a hand enters the designated location and is detected. Torsor-distance Meter measures the distance from the character's torsor to the wrist, and if the value is below a certain value, it is judged as an action step and the current time is recorded. Using these two modules, the expected motion time of the sign language word motion data is specified, and the motion data is modified according to the time. The edited data is stored in the DB so that the motion can be seen naturally when Korean sign language animation is created later.

### 4.1.2 Hand Motion Combine (HMC) Algorithm

Sign language can express new meanings by combining the movements of the two arms in various ways. In other words, when emphasizing instructions such as ‘this one’, ‘this way’, or indicating the size of an object by one hand, two sign language are combined. Since words representing indications or emphasis are independently defined in sign language dictionaries, an algorithm that synthesizes two independent sign words into a single sign word is required for a new expression. That is, the sign word motion corresponding to the left-hand motion and the sign word motion corresponding to the right-hand motion are synthesized into one independent word, and the process proceeds through the algorithm shown in Figure 5.

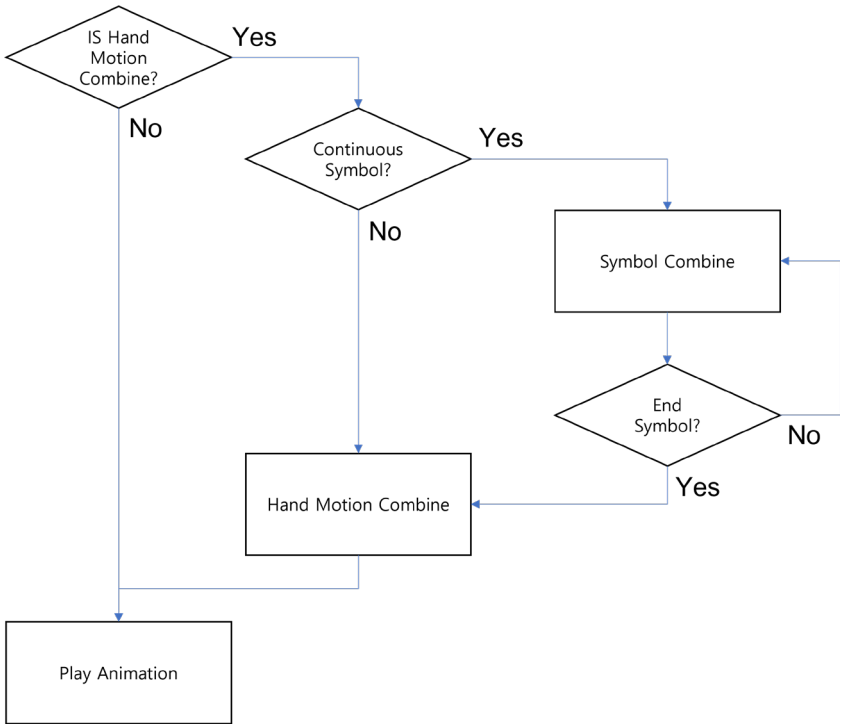


**Fig. 5.** Automatic Sign Language generation Algorithm diagram. The algorithm determines whether a Hand Motion Combination is necessary and, if necessary, determines whether a word is continuous symbol. If it is continuous, do the symbol combination. After then, it is used for animation playback after Hand Motion Combination.

First, the entire input data is scanned to review whether synthesis of left-hand and right-hand motions is necessary. As a result of the sentence scan, if it is determined that the synthesis of the left-right hand motion is necessary, it is determined whether the sign language word motion corresponding to the right-hand motion is a continuous motion consisting of an explanatory word. In this process, if it is not a continuous motion composed of language words, the two motions are synthesized through the HMC.

### 4.1.3 Symbol combine algorithm

In the automatic sign language animation generation algorithm, animation is created with data of one word unit. JIHWA-words are composed of phoneme units of characters such as 'ㅇ' and 'ㅣ'. In the algorithm, one phoneme corresponds to one word, so in the case of the continuous action of the right hand JIWHA-words, one left hand action will be combined with the right-hand actions. In this case, since the left-hand motion is repeated whenever the right hand motion is changed, the animation

becomes unnatural, so the right hand motion must be combined into one motion and then combined with the left hand motion.

In the HMC synthesis step, if the sign word motion corresponding to the right-hand motion is a continuous motion consisting of JIWHA-word, a process of combining the continuous motion into one motion data (this is called Symbol Combine) is executed. Symbol Combine uses Unity's additional modules Recorder Module and FBX Recorder Module. [8][9] Whenever one JIWHA-word is combined, the algorithm determines whether the continuous motion has been completed, and all JIWHA-words are combined to form one sign language motion. After making it, it is synthesized into a new word by applying the HMC algorithm.

## 4.2 Animation play procedure

### 4.2.1 Encoding Decoding Algorithm

In the previous paper, sign language animation was generated through the encoding-decoding algorithm [10]. The encoding-decoding algorithm is an algorithm that analyzes and encodes input data, then decodes and animates the encoded data in a generator. The encoding and decoding process is performed as shown in Figure 6. In the encoding process, the input data is analyzed through the DB, synthesized into word descriptors, and sentence data composed of word descriptors is sent to the generator. In the decoding process, sentence data is decoded in units of word descriptors, and playback variables in word data are analyzed, and then animations are displayed at appropriate timings. Through this, a user may display a sign language animation by combining words for which data exists without complicated manipulation.

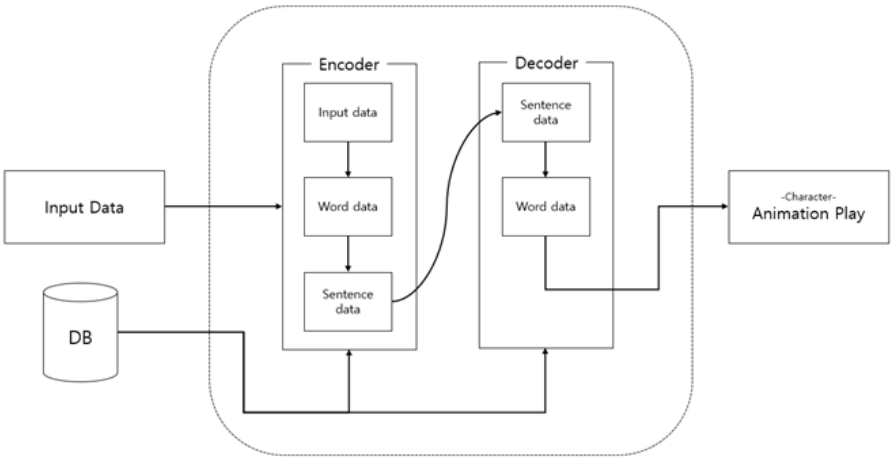


**Fig. 6.** Encoding Decoding Algorithm diagram. Input data encode to Sentence data in encoder. After then, decoder receive the Sentence data and decode that data to Word data. In Word data, there were various variables to control animation play.

## 5 Animation result

To confirm the results of the algorithm presented in this paper, motion capture data used in previous studies were used. [11] Figure 7 shows the result of HMC when the left-hand words are not continuous JIHWA-words. Figure 8 shows the result of HMC when the left-hand words are continuous JIHWA-words, and in this case, Symbol Combine is used to make the JIHWA-words into one word. Figure 9 shows the result of HMC using the motion words created in Figure 8. The sentence playback result is shown in Figure 10. The motion used at points 8 and 9 in Figure 10 is the same HMC + Symbol result as in Figure 9, and through this, the connection between the motions can be seen to be natural.

Furthermore, when the automatically generated Korean sign language animation was applied through the demonstration service conducted in joint research, it was confirmed that the understanding increased. This indicates that the animation is playing an important role in improving the senses and understanding of users and providing a better experience. Figure 11 shows the photos and app screens at the time of the demonstration service. Figure 12 shows the change in satisfaction after the demonstration service, and it indicates that the overall satisfaction increased when the animation was applied.

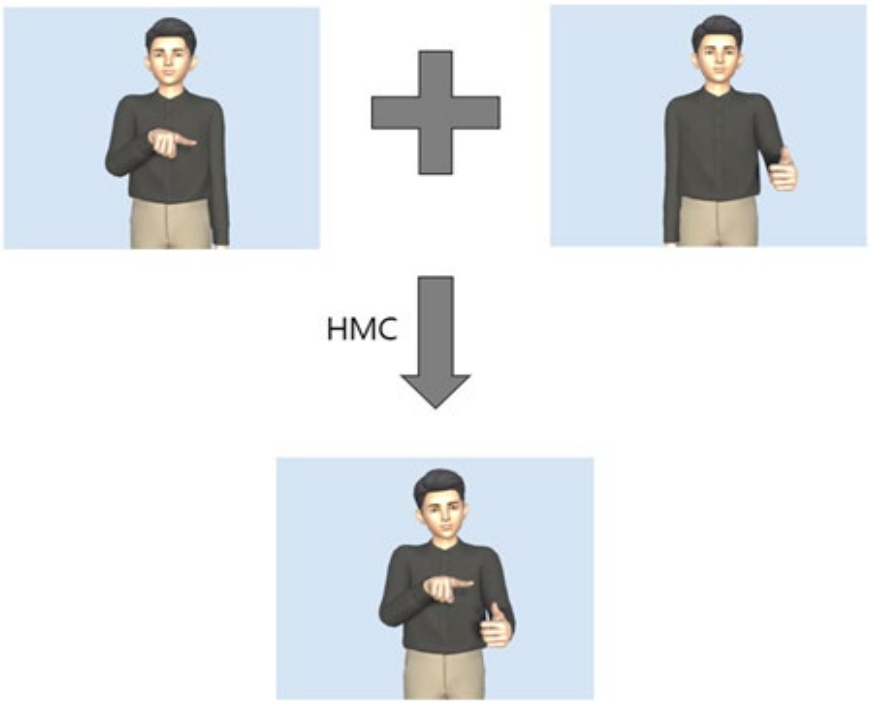


**Fig. 7.** Output of Hand Motion Combine Algorithm. It can be seen the left-hand motion and the right-hand motion are combined to become one new motion.

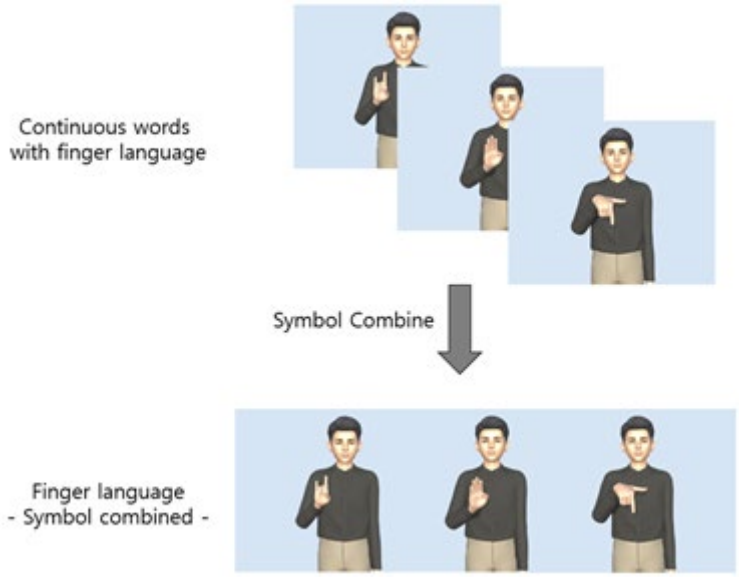


**Fig. 8.** Output of Symbol Combine Algorithm. Continuous words with finger language are combined to one right-hand motion.

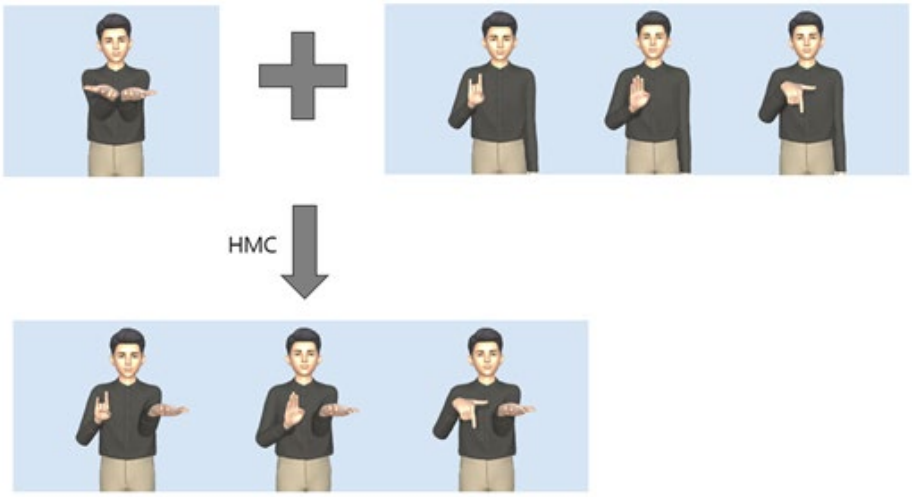


**Fig. 9.** Output of Hand Motion Combine Algorithm after Symbol Combine. The result of Figure 8 is combined on the right-hand motion. The right-hand motion (Continuous finger language words) is reproduced while maintaining the left-hand motion

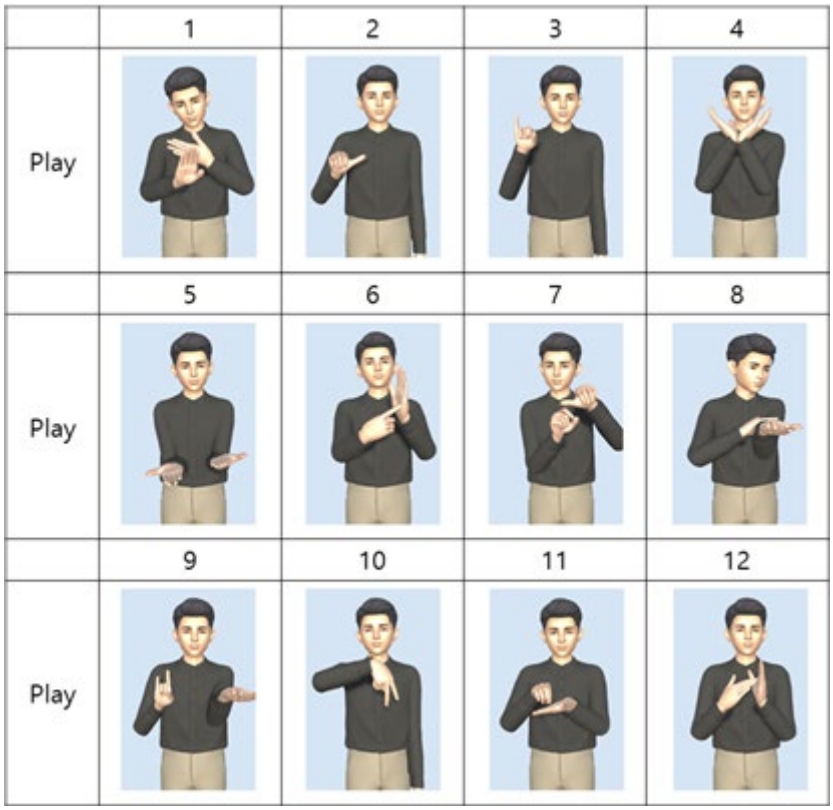


**Fig. 10.** Output of playing Sign Language Sentence. It is a reproduced sentence including the Figure 9. In frame 9, The HMC worked fine.


**Fig. 11.** Left shows the photo at the time of the demonstration service. Right shows the app screen at the time of the demonstration service.

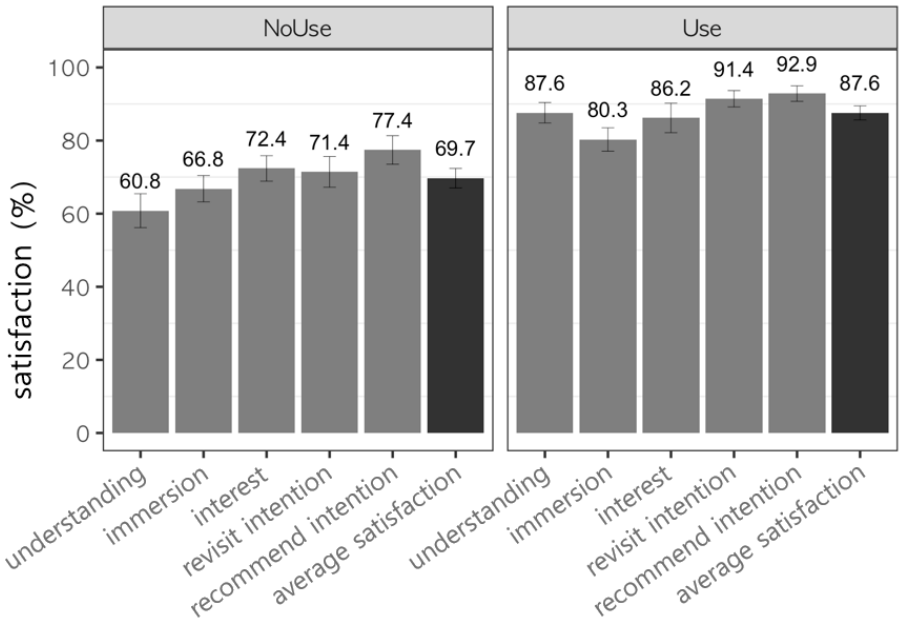


**Fig. 12.** The change in satisfaction before and after using application. It can be seen overall satisfaction has increased after using it.

## 6 Conclusion

Korean sign language is a language with a different gramma system from Korean, and technology to automatically convert Korean into Korean sign language animation is needed for the smooth cultural life of the hearing-impaired. To do this, a Korean-sign language conversion dictionary is required. If the sign language dictionary has sign language motion data corresponding to sign language words, it is possible to express a Korean sentence in sign language by continuously connecting sign language words. However, if the sign language words stored in the dictionary are connected as they are, an unnatural motion occurs, and the meaning is not properly conveyed. Since the meaning of a sign language is determined by its shape of hands, and a sign language with a new meaning can be created through a combination of two sign language actions, an innumerable number of words that can be created by combining sign words must be captured in a sign language dictionary. However, it is realistically impossible to capture an innumerable number of sign language words. To solve these problems, this paper describes a system that reproduces Korean sentences as sign language sentences after preprocessing sign language words. Motions including start and end motions were refined through pre-processing, and when a word composed of a combination of two sign language motions was needed, a synthesis algorithm was

used to synthesize them, and then sign language animation was reproduced using an encoding-decoding algorithm.

In the future, we plan to improve the completeness of the system by adding basic words necessary for motion synthesis and correcting minute discontinuities that appear after motion synthesis.

**Acknowledgements**
This research is supported by Ministry of Culture, Sports, and Tourism (MCST) and Korea Creative Agency (KOCCA) in the Culture Technology (CT) Research & Development Program (R2020060002) 2022.

## References

1. Mi Jeong Lee, Sun Hwa Lee, Jin Hwe Kim, “A Survey on the Utilization of Welfare Services for the hearing-impaired and Their Needs”, Korea Disabled People Development Institute, 2010.
2. Seong Ok Won, et al. “Korean Sign Language and Grammar Research”, Korea Welfare University Industry-Academic Cooperation Group, 2019.
3. Introduction of Korean Sign language, https://www.korean.go.kr/front/page/pageView.do?page_id=P000300&mn_id=202
4. Vu, Dang Thanh, et al. "Text Data Augmentation for the Korean Language." *Applied Sciences,* 12.7 (2022).
5. Young Min Ko, et al. “A Sign Word Description Model for Efficient Generation of Sign Language”, Korea Smart media Conference, 10(1), pp.15-18, 2021.
6. Young Min Ko, et al. "The Development of Animation Technology for Intelligent Sign Language Transformation for the Hearing Impaired", MITA, 2020.
7. Jong Ho Jeong, et al. “A FBX File Combination Algorithm for Natural Sign Language Animation”, Korea Smart media Conference, 2021.
8. Unity Recorder module, https://docs.unity3d.com/Packages/com.unity.recorder@2.2/manual/index.html
9. Unity FBX Recorder module, https://docs.unity3d.com/Packages/com.unity.formats.fbx@4.0/manual/recorder.html
10. Jong Ho Jeong, et al. “Advanced Sign Word Generation Algorithm for Natural Animation”, Korea Smart media Conference, 2022.
11. Jong Woo Seo, et al. “A study on Korean sign language translation script method”, a Study on the Integrated Academic Presentation of the Korean Management Association, 191-198, 2021