

AVATAR-BASED SIGN LANGUAGE INTERPRETATION FOR WEATHER FORECAST AND OTHER TV PROGRAMS

Juhyun Oh¹, Byungsun Kim¹, Minho Kim², Sangwook Kang², Hyuckchul Kwon², Iktae Kim³ and Youngho Song³

¹Korean Broadcasting System (KBS), Korea, ²Pusan National University (PNU), Korea and ³Bluenine Inc., Korea

ABSTRACT

Although the closed caption broadcasting is provided for almost the whole broadcast time in Korea, the sign language interpretation is not. By translating the closed captions, it is possible to provide three-dimensional (3D) sign language translations for more broadcast time. We propose a sign language broadcasting system for weather forecasts, and extend it for all kind of TV programs. In order to find the frequency of each word, we analysed the last three years' weather forecast scripts, and an open-domain corpus of about 1.2 million words from KBS. We use the Korean wordnet, *KorLex*, to build the sign language synonym dictionary, and for the word sense disambiguation to improve the translation performance. Optically-captured sign language motions are used for the 3D avatar to present sign language with motion blending. We developed a mobile on-demand sign language weather forecast application, and a real-time sign language interpretation system for all kind of TV programs.

INTRODUCTION

Closed caption and sign language broadcasting are provided with the terrestrial digital television (DTV) services for deaf people in most countries. Closed caption broadcasting can be 'hidden' by the DTV standard (1) and does not disturb the hearing (non-disabled) viewers' watching, and therefore can be serviced for almost the whole broadcast time. However, the sign language interpretation, aside from the cost problem, occupies some space in the TV screen and is hardly provided for more than 5% of the broadcast time in Korea. Even if a TV program has a sign language interpretation in it, deaf viewers are still not satisfied because it is small and fixed in the screen.


If one can generate the sign language interpretation from the closed captions, the sign language interpretation can be provided for the remaining 95% of the broadcast time. To this end, the real-time Korean-Korean sign language (KSL) translation and 3D avatar animation technologies are required. Automatic 3D avatar-based interpretation systems have been proposed for a few sign languages. For example, Kaneko et al (2) suggested a Japanese sign language (JSL) interpretation system using TV program making language (TVML). Araújo et al (3) proposed an automatic translation and middleware structure for DTV for Brazilian sign language (LIBRAS).

We propose a system that translates the closed captions of weather forecast programs into KSL and present it with 3D avatar animation. The translated sign language data are sent using the Internet to the receivers such as a personal computer (PC) and mobile devices to show the corresponding sign language animation. The system consists of the Korean-KSL translator, the sign language avatar animation system, and the server system that provides the closed caption and video for the most recent weather forecast in an on-demand manner. We extend the translation and the animation systems for all kind of TV programs other than the weather forecast. Similarly to the existing systems, the purpose of the proposed system in this paper is to provide information even when human interpreters are not available, not to replace them.

CLOSED CAPTION TO SIGN LANGUAGE TRANSLATION

Table 1 shows an example of Korean-KSL translation in a couple of weather forecast sentences by a professional sign language interpreter. Our goal is to implement an automatic translator that outputs similar KSL results when the same Korean weather forecast input is given.

Table 1 – An example of manual Korean-KSL translation in weather forecast scripts (expressed in English for readability).

Korean	Recently it is cold at every weekend.
KSL	week + end + cold 
Korean	It is raining from the early morning in the southern area and Jeju Island.
KSL	warm + place + and + Jeju Island + place + morning + from + rain

SIGN LANGUAGE DICTIONARY

There are about 12,000 words in KSL Dictionary. It is difficult to build the Korean-KSL dictionary and capture the motions for all these words. Therefore we analysed the weather forecast scripts for the past three years, from KBS and a few other sources. After some preprocessing of the weather forecast scripts, the scripts are divided into part of speech (POS). For this purpose we use the POS tagger of PNU with about 1 million registered words (4). The accuracy of the POS tagger is about 98%. Since the basic word orders of Korean and KSL are similar as subject-object-verb (S-O-V), direct word-to-word

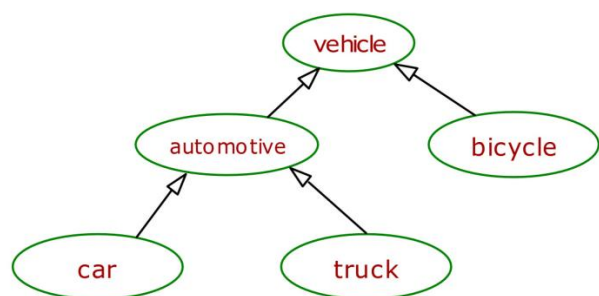


Figure 1 – A simple Wordnet example

translation rule is applied in the proposed system.

Synonyms are translated into same sign language words. For example, ‘house’, ‘housing’ and ‘abode’ are synonyms and all should be translated into one sign language word. The synonym dictionary is built based on *KorLex* by Yoon et al (5), which is the Korean Wordnet. Using this synonym dictionary, a word that is absent in the KSL dictionary still can be translated to a synonym, increasing the translation success rate. In Figure 1, if the word ‘truck’ is not registered in the dictionary, the word still can be translated to ‘car’ or ‘vehicle’. Without the synonym dictionary, non-registered words would have been omitted or represented with finger spelling, making the system less useful.

WORD SENSE DISAMBIGUATION (WSD) WITH KORLEX

Words that have multiple meanings have ambiguities and the proper meaning in the context must be found for correct translation. For example, the same word ‘sa-gwa’ in Korean is used to represent two different meanings, ‘apology’ and ‘apple’.

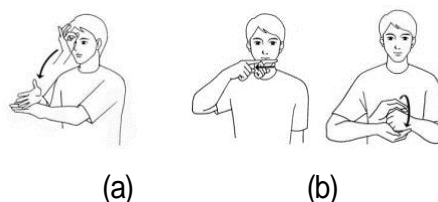


Figure 2 – Sign language for ‘sa-gwa’.

(a) sa-gwa¹ (apology) (b) sa-gwa² (apple)

To determine the meaning of a word, context words around it are considered. If the context words such as ‘formal’ or ‘announcement’ appear together with the word ‘sa-gwa’, it is translated to sa-gwa¹ (apology). If the context words ‘eat’ or ‘juice’ appears, sa-gwa² (apple) is chosen. In addition to the word ‘sa-gwa’, the related words (for example, siblings in *KorLex*) such as ‘orange’, ‘peach’, ‘gratitude’, and ‘appreciation’ are used together to compute the χ^2 values to increase the recall, as shown in Table 2.

Table 2 – The siblings of ‘sa-gwa’ and χ^2 values with the context words

Words appearing together with ‘sa-gwa’	χ^2 with sa-gwa ¹ (apology)	χ^2 with sa-gwa ² (apple)
‘eat’ (verb)	1.25	86.94
‘juice’ (verb)	0.01	17958.63
‘government’ (noun)	3.88	0.01
‘people’ (noun)	9.23	0.01
‘formal’ (adj.)	70.05	0.01
‘announcement’ (noun)	25392.88	108.05

3D AVATAR ANIMATION

In order to visualize the translated sign language codes, sign language motion database must have been built. One can make the motions by editing each key frame, but we capture the human interpreter’s motions because they are much more natural. In addition to the existing 2,700 sign language motions for daily life, we captured 500 words dedicated for weather forecast at a studio. Afterwards, we added 1,000 words for the open-domain

interpretation. 15 *Vicon* motion capture cameras and 41 infra-red markers were used to capture the body motion of a professional sign language interpreter, as shown in Figure 3. *CyberGlove* was used to capture the hand and finger motion.

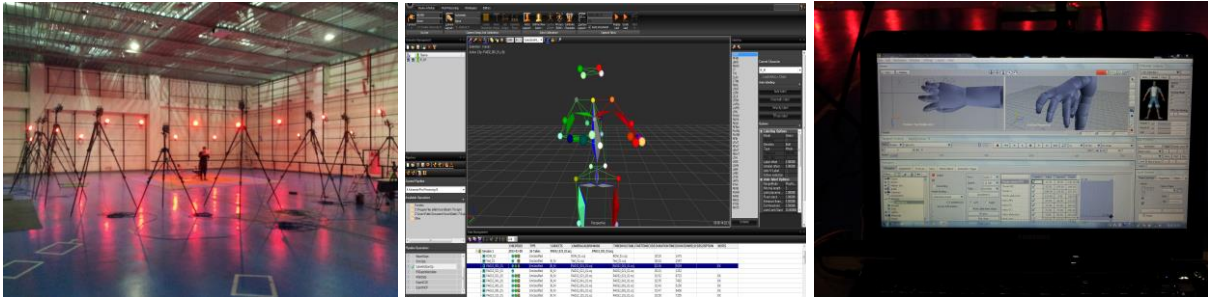


Figure 3 – The motion capture studio and the captured body and hand motions

Body and hand motions are combined together in the *MotionBuilder* software. We manually filter the noisy motion capture data. The original motion data at 250 frames per second (fps) are resampled to 30 fps. Marker interferences are resolved at this time. The *CyberGlove* captures the right hand and fingers' motion at 90 samples per second. The left hand motion is copied from the right one or edited manually. Figure 4 shows the body and hand motions combined together in the *MotionBuilder* software.

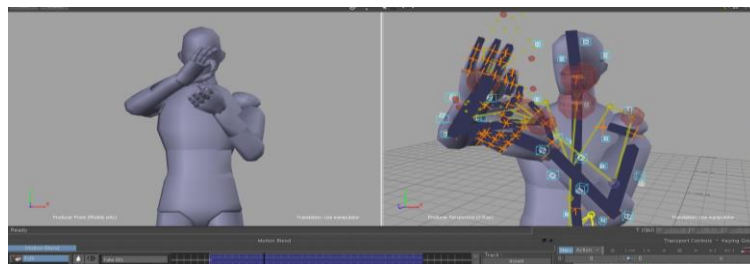


Figure 4 – The combined motion

We built a new 3D avatar for sign language animation using 8,000 polygons. Figure 5 shows the avatar model, and the motions applied to the avatar.



Figure 5 – The avatar model and the motions applied to it

RESULTS OF THE PROPOSED SYSTEM

Figure 6 shows the results of the proposed system, for weather forecast and the other real-time TV programs, in the forms of mobile applications and Windows-based software. For the multi-platform applications, we used the *Unity3D* engine.



Figure 6 – Demonstration of the proposed system. (a) Weather forecast application, (b) mobile application for real-time sign language interpretation, (c) PC software playing video and script files at offline mode, (d) the output of the PC software (online mode) composited on the live TV screen using a chromakeyer.

The user can move and resize the avatar, and change the color of the clothes as shown in Figure 6. The user can also change the background color for chromakeying (only for demonstration purpose).

Table 3 shows the translation rate for the weather forecast, for the internal and external corpus. It is shown that the translation success rate is improved using *KorLex*. Weather forecast is a relatively easy domain in the machine translation field. The translation success rate decreases when the domain is open. Table 4 shows the translation rate results for small-sized corpus of a few different TV program categories.

Table 3 – Translation results of weather forecast

Corpus type	Corpus size (words)	Word translation rate	
		without <i>KorLex</i>	with <i>KorLex</i>
Internal corpus	82,303	95.87%	96.17%
External corpus	1,448	88.60%	90.68%

Table 4 - Open-domain translation results

Category	Corpus size (words)	Word translation rate
News	250	61%
Sports	242	61%
Drama	238	67%

CONCLUSIONS

The system proposed in this paper generates 3D sign language animation by translating the closed captions in DTV, for the deaf people to view the weather forecast with sign language. In order to find the frequency of each word, we analysed the last three years' weather forecast scripts from several sources. We built sign language synonym dictionary using *KorLex*, to improve the translation performance. *KorLex* was also used for the word sense disambiguation process. We optically captured the motions of a professional sign language interpreter and built the motion database. The motions were applied to a 3D avatar with motion blending. We extended the sign language weather forecast system's dictionaries, translation rules, and motion database, for the real-time open-domain sign language interpretation system. We consider using the proposed real-time sign language interpretation as an auxiliary channel for N-screen services such as 'my K' (6).

Although the proposed system is targeted only for KSL users, we believe that it will be helpful to the other similar researches, especially if the source language and target sign language have a similar word order as in Korean and KSL. In the future work, we will attempt to develop a statistical translation system to improve the translation rate, and build large-scale Korean-KSL translation corpus.

REFERENCES

1. CEA: "Digital Television (DTV) Closed Captioning," Doc. CEA-708-D, Consumer Electronics Association, Arlington, VA, August.
2. H. Kaneko, N. Hamaguchi, M. Doke and S. Inoue, 2010. Sign language animation using TVML. Proceedings of the 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry. ACM.



3. T. Araújo, F. Ferreira, D. Silva, F. Lemos, G. Neto, D. Omaia, G. Filho, T. Tavares. 2012. Automatic generation of Brazilian sign language windows for digital TV systems, Journal of the Brazilian Computer Society 19.2 107-125.
4. <http://klpl.re.pusan.ac.kr/>, 2015.
5. A. Yoon, S. Hwang, E. Lee, H. Kwon. 2009. Construction of Korean Wordnet KorLex 1.5. Journal of KIISE: Software and Applications, 36(1), 92-108 (in Korean).
6. <http://k.kbs.co.kr/>, 2015.