An Avatar-Based Weather Forecast Sign Language System for the Hearing-Impaired

Juhyun Oh, Seonggyu Jeon, Minho Kim, Hyukchul Kwon, Iktae Kim

HAL Id: hal-01391353
https://inria.hal.science/hal-01391353

Submitted on 3 Nov 2016

An Avatar-Based Weather Forecast Sign Language System for the Hearing-Impaired

Juhyun Oh[1], Seonggyu Jeon[1], Minho Kim[2], Hyukchul Kwon[2], and Iktae Kim[3]

[1]Technical Research Institute, Korean Broadcasting System, Seoul, Korea
{jhoh,cheez}@kbs.co.kr
[2]Dept. of Computer Science, Pusan National University, Busan, Korea
{karma,hckwon}@pnu.ac.kr
[3]Bluenine Inc., Seoul, Korea
blue9naa@naver.com

Abstract. In this paper, we describe a text-to-animation framework for TV weather forecast sign language presentation. To this end, we analyzed the last three years' weather forecast scripts to obtain the frequency of each word and determine the order of motion capture. About 500 sign language words were chosen and motion-captured for the weather forecast purpose, in addition to the existing 2,700 motions prebuilt for daily life. Words that are absent in the sign language dictionary are replaced with synonyms registered in *KorLex*, the Korean Wordnet, to improve the translation performance. The weather forecast with sign language is serviced via the Internet in an on-demand manner and can be viewed by PC or mobile devices.

Keywords: sign language, avatar animation, closed caption, machine translation, wordnet

1 Introduction

Closed caption and sign language broadcasts are provided with the terrestrial digital television (DTV) services in Korea for the hearing-impaired people. Even though closed captions are provided for almost the whole broadcast time, sign language broadcasting covers only about 5% of it. We propose a system that translates the closed captions of weather forecast programs into Korean sign language (KSL) and present it with three-dimensional (3D) avatar animation. The translated sign language data are sent via the Internet for the receivers such as 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.

A similar system for Japanese sign language (JSL) was proposed to use TV program making language (TVML) [1,2] for high-quality computer animation. However the translation is not automatic in their work and the user should input the sign language words in the order of JSL. For Brazilian sign language (LIBRAS), automatic

translation and middleware structure for DTV was also proposed in [3,4]. Similarly to the existing studies, the purpose of the proposed system in this paper is to provide information when human interpreters are not available and not to replace human interpreters.

This paper is organized as follows. In Section 2, the Korean-KSL translation method is described. In Section 3, the 3D avatar animation scheme is introduced. In Section 4, the system to provide the weather forecast sign language service is described. The experimental results and conclusion are given in Section 5 and 6 respectively.

2 Korean-KSL 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.

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

2.1 Korean-KSL Dictionary

There are about 10,000 words in KSL Dictionary [5]. It is very difficult to build the Korean-KSL dictionary and capture the motions for all these 10,000 words. Thus we investigated the weather forecast scripts for the past three years, from Korean Broadcasting System (KBS) and a few other sources. After some preprocessing of the weather forecast scripts, the scripts are divided by part of speech (POS). For this purpose we use the POS tagger of Pusan National University with about 1 million registered words [6]. 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 translation rule is applied in the proposed system.

Table 2 shows ten most frequent words for noun, adjective, verb, and adverb in the analyzed scripts. Notice that the POS in Korean is different from that of English for some words. In Table 2, the most frequently used noun and adjective are 'temperature' and 'high' respectively. The cumulative frequency in Table 2 is important because it shows the translation capability when that word and the above ones in the table were registered in the dictionary. For example, if only ten most frequently used nouns were registered in the dictionary and motion-captured, the system can probabilistically process 33% of the nouns in the script.

Table 2. Ten most frequently used noun, adjective, verb, adverb words (expressed in English) in the weather forecast scripts. C.F. means the cumulative frequency of each POS.

Noun			Adjective			Verb			Adverb		
Word (Kor/Eng)	Frequency (times)	C.F. (%)	Word	Freq.	C.F.	Word	Freq.	C.F.	Word	Freq.	C.F.
temperature	4,180	6	high	1,277	14	go down	1,451	11	a little	742	11
tomorrow	3,295	11	many	1,173	27	rise	854	18	gradually	702	21
day	2,664	15	exist	1,125	40	come	844	24	mostly	628	30
today	2,602	19	clear	841	49	rise	821	31	again	576	38
rain	2,092	22	low	798	58	continue	737	36	somewhat	470	45
Seoul	1,747	24	strong	609	65	hang	712	42	more	417	51
morning	1,611	27	similar	480	70	take	628	46	occasionally	284	55
region	1,585	29	large	376	74	pass by	493	50	again	269	59
central area	1,426	31	dark	343	78	fall	440	54	especially	204	61
whole country	1,293	33	cold	313	81	be seen	430	57	but	185	64

2.2 KSL Synonym Dictionary

Synonyms are translated into same sign language words. For example, 'house', 'housing' and 'abode' are synonyms and all translated into one sign language word. The synonym dictionary is built based on KorLex [7], as shown in Fig.1. Using this synonym dictionary, a word that is absent in the KSL dictionary can be translated to a synonym, increasing the translation success rate. Without the synonym dictionary, non-registered words would have been represented with finger spelling, decreasing the usefulness of the system.
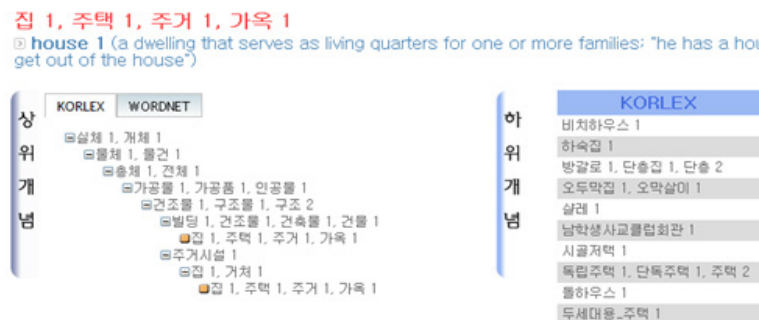

Fig. 1. The word '집(house)' found in KorLex.

2.3 Word Sense Disambiguation with KorLex

In Korean, the word '눈' (pronounced as 'nun') can mean 'eye' or 'snow'. This kind of ambiguity should be resolved in the context. A simple rule to determine the

word '눈' as 'snow' is shown in Table 3. If the context words (CW) exists around the ambiguous word, then it is translated to 'snow'.

Table 3. A simple rule for the ambiguous word '눈'

눈 CW → snow	
(R1)	CW = [내리다(fall)	날리다(fly)	흩날리다(flutter)]

This kind of rules suppose the exact match to the CWs, therefore its recall is low. We replace the CWs with the synonym set number in KorLex, to increase the recall. Case particle restriction (R3) and CW verb conjugation restriction (R4) are also applied, as shown in Table 4.

Table 4. Modified rule for the ambiguous word '눈'

눈+P CW → snow	
(R1)	Context = B1
(R2)	CW = [02674938	02041026	02197925]
(R3)	P = [subjective, complement, objective, auxiliary particle]
(R4)	Conjugation = 1001 + 2001

3   Sign Language Avatar Animation

We use about 2,700 sign language motions for daily life owned by Primpo Inc [8]. Additionally, we captured 507 words dedicated for weather forecast at a studio. 15 Vicon motion capture cameras [9] and 41 infra-red markers were used to capture the motion of a professional sign language interpreter, as shown in Fig.2. Cyberglove [10] was used to capture the hand motion.
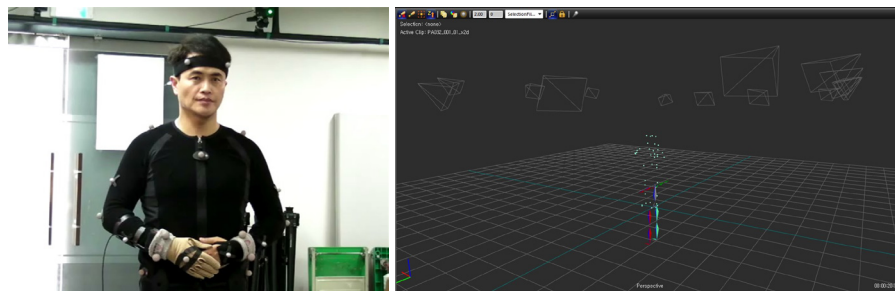

Fig. 2. The sign language interpreter (left) and the corresponding captured motion (right).

A 3D female avatar model was built for sign language animation. Fig.3 shows the avatar model and bones associated to it.
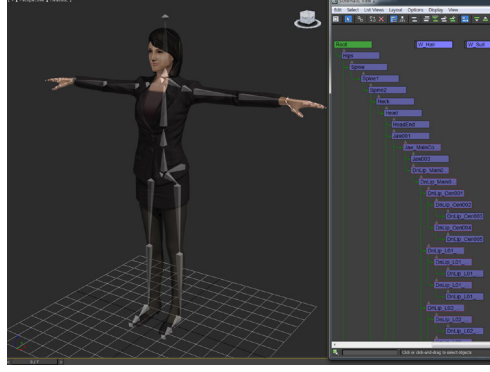
Fig. 3. 3D avatar and bones.

When animating a sign language sentence, two contiguous sign words are over-lapped to each other similarly in human sign language. 15 frames at the end of each sign word are blended with those of the following word, using a linear motion blend-ing equation [11]. Let two sign word motions to blend be $\mathbf{M}_i$ and $\mathbf{M}_j$, and the k-th joint of concern $\mathbf{J}_{ik}$ and $\mathbf{J}_{jk}$. If the two motions are blended with the weights w_i and w_j respectively, the blended position of the k-th joint is given as

$$\mathbf{J}_{k_{blended}} = \frac{\mathbf{J}_{i_k} \cdot w_i + \mathbf{J}_{j_k} \cdot w_j}{w_i + w_j} \tag{1}$$

where $0 \le w_i \le 1$ and $w_j = 1 - w_i$ in this case. Similarly the blended quaternion $q_{blended}$ is given as

$$q_{blended} = q_i \left(q_i^{-1} q_j \right)^{w_j} \tag{2}$$

where q_i and q_j represent the quaternions to be blended [12]. Fig.4 shows a blended motion frame using the original motions of two different sign words.
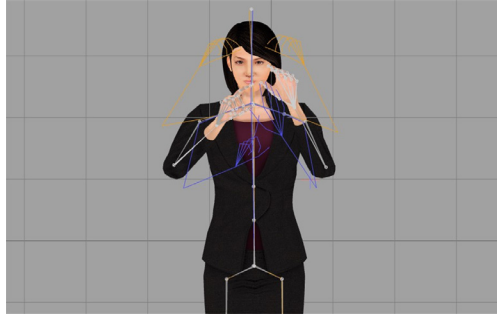

Fig. 4. Motion blending.

4   The Weather Forecast Sign Language System

In the service scenario, when a hearing-impaired user launches the 'weather forecast' application on a PC or mobile devices, the latest weather forecast video is provided with sign language animation. Fig.5 shows the system for the weather forecast sign language service.
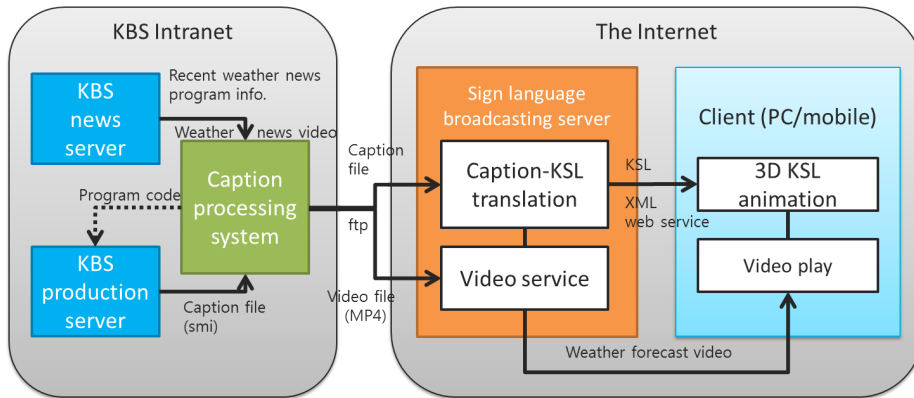

Fig. 5. Weather forecast sign language system

Fig. 6. Simplified XML schema for the KSL translation result

The 'caption processing system' asks the news server for the latest weather forecast information. If there is a new weather forecast, it inquires the corresponding closed caption from the production server, using the program code of the weather forecast. The weather forecast video in Flash Video (FLV) format originally targeted for real-time message protocol (RTMP) streaming is transcoded to H.264 (MP4). The closed caption is translated into KSL with the method described in section 2 in the sign language broadcasting server. The closed caption in KSL is distributed to clients using an XML web service. In a client device, weather forecast video and sign language animation are composited together and presented to the hearing-impaired user. Fig.6 shows the simplified XML schema for the KSL translation result.

5   Experimental Results

Fig.7 shows the result screens on a PC and an Android phone. We use Unity3D engine [13] to develop an application for both PC and mobile devices. The avatar can be moved to left or right, and can be enlarged or reduced, as the hearing-impaired users want the sign language broadcasting to be.

Fig. 7. Sign language animation on a PC (top) and an Android (bottom) phone.

To our best knowledge, there has been no similar research that translates weather forecast texts into KSL to generate sign language animation. Hence we quantitatively compare the translation results in Table 5, with and without the Korlex-based synonym dictionary and word sense disambiguation.

Table 5. Quantatative comparison of the word translation rate.

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

In Table 5 the word translation rate is given as the ratio of the number of the words for which correct translation words were chosen, to the number of total words in the corpus. By applying KorLex the translation rates were increased, especially in terms of the translation failures. For example, the translation failure rate for the external corpus was reduced from 11.4% to 9.32% by applying KorLex. Notice that this is only a simplified measure for the evaluation purpose, because the characteristics of sign language such as non-manual signals and spatial representation are difficult to include in the evaluation, and were not considered in this paper.

6   Conclusion

The system proposed in this paper generates 3D sign language animation by translating the closed captions in DTV, for the hearing-impaired people to view the weather forecast with sign language. In order to find the frequency of each word, we analyzed 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 captured the motions of the additional 500 words for the weather forecast purpose. The motions were applied to a 3D avatar with motion blending. The servers and applications were developed to show weather forecast video and avatar animation in an on-demand manner.

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 expand the translator and motion database for all kind of TV programs other than the weather forecast, and develop a real-time sign language broadcasting system. Constructing a sign language corpus using daily TV programs can be another future work.

References

1. H. Kaneko, N. Hamaguchi, M. Doke, S. Inoue. Sign language animation using TVML, Proceedings of the 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry. ACM. (2010)
2. N. Kato, H. Kaneko, S. Inoue, T. Shimizu, N. Hiruma. Machine translation to Sign Language with CG-animation. ABU Technical Review 245. (2011)

3. T. Araújo, F. Ferreira, D. Silva, F. Lemos, G. Neto, D. Omaia, G. Filho, T. Tavares. Automatic generation of Brazilian sign language windows for digital TV systems, Journal of the Brazilian Computer Society 19.2 107-125. (2013)

4. F. Ferreira, F. Lemos, G. Neto, T. Araújo, G. Filho. Providing support for sign language in middlewares compliant with ITU J.202. IEEE International Symposium on Multimedia. (2011)

5. http://www.korean.go.kr/. (2014)

6. M. Kang. Variable-Length Category-Based N-Gram Part-of-Speech Tagging Model Considering Morph-Syntactic Characteristics of Korean, Doctoral thesis, Pusan National University, (in Korean). (2007)

7. A. Yoon, S. Hwang, E. Lee, H. Kwon. Construction of Korean Wordnet KorLex 1.5. Journal of KIISE: Software and Applications, 36(1), 92-108 (in Korean). (2009)

8. http://www.primpo.co.kr/. (2014)

9. http://www.xsens.com/products/. (2014)

10. CyberGlove Systems. CyberGlove II brochure, http://www.cyberglovesystems.com/products/cyberglove-II/overview. (2009)

11. K. Slot. Motion Blending. Copenhagen University. Department of Computer Science. (2007)

12. K. Shoemake. Animating rotation with quaternion curves. Proceedings of ACM SIGGRAPH 85. Annual Conference Series. 245-254. ACM Press. (1985)

13. https://unity3d.com/unity/. (2014)