

한국 수화 아바타를 위한 수화 사전의 개선 방법

오영준¹, 박광현², 변증남³
한국과학기술원 인간친화 복지 로봇시스템 연구센터¹
한국과학기술원 전자전산학과²³
{yjooh¹, akaii²}@ctrsys.kaist.ac.kr
zbien@ee.kaist.ac.kr³

Improvement of Sign Word Dictionary for Korean Sign Language Avatar

Young-Joon Oh¹, Kwang-Hyun Park², Zeungnam Bien³
Human-friendly Welfare Robot System Engineering Research Center, KAIST¹
Dept. of Electrical Engineering and Computer Science, KAIST²³

요 약

본 논문에서는 수화 아바타가 실제 청각장애인처럼 자연스러운 수화 동작을 표현하면서 정확한 의사를 전달할 수 있도록 동음이의어에 대한 처리를 다룬다. 기존의 수화 사전에 품사 정보를 추가하고 한글 형태소 분석기를 활용하여 동음이의어를 구분할 수 있도록 수화 사전을 개선하는 방법을 제안한다.

1. 서 론

1.1 수화

수화는 청각 장애인의 중요한 의사소통 수단으로서 음성 언어와 대등한 수준의 언어이며[1] 손가락 동작과 비손가락 동작(NMS, Non-Manual Signals)을 동시에 사용한다. 또한, 손짓과 함께 동작해야 할 신체 움직임, 얼굴 표정 등으로 구성되는 의사소통 체계를 가지고 있으며, 생각과 감정을 정확하게 표현할 수 있는 몸짓 언어이다[2]. 일상생활에서 의사소통을 하고 원만한 인간 관계를 유지하게 하는 수화는 청각장애인의 삶에서 큰 비중을 차지한다. 최근에 인터넷 기술의 발전으로 청각 장애인들이 화상채팅 메신저와 영상전화, 화상 휴대전화, 전화중계 서비스를 이용하는 빈도가 증가하고 있지만, 건청인과 청각장애인 사이에 수화통역을 하는 시스템이 없어 불편함이 있다.

1.2 수화와 관련된 기존 연구

수화에 기반하여 의사소통을 돕는 시스템으로는 수화 인식 시스템과 수화 발생/표현 시스템이 있는데, 본 논문에서는 수화 발생 시스템을 다룬다. 기존 연구로는 한국과학기술원(KAIST) 변증남 교수 연구실에서 1995년 초에 한글 문장을 수화 동작 그래픽으로 보여 주는 '한글 표준수화 자동발생 시스템'을 개발하였다[3]. 2000년에는 PC용 자동 수화 발생 시스템을 개발하였으며[4], 2001년에는 한글 형태소 분석 알고리즘을 도입하

여 TV 자막 방송의 수화 발생 시스템을 개발하였다[5]. 또한, 2007년 3월에는 작업현장에서 건청인이 사용하는 문장을 청각장애인의 수화 문장으로 번역하여 아바타의 수화 동작으로 보여주는 '한국 수화 메시지 보드 시스템'을 개발하였다[6].

일반 문장을 수화 문장으로 번역할 때에는 한글 형태소 분석기가 분석한 결과에서 청각장애인이 사용하지 않는 조사, 어미, 접사, 존칭, 시제 등은 생략하는데[5], 기존의 수화 발생 시스템은 한글 형태소 분석 과정에서 품사가 다름에도 불구하고 동음이의어를 구분하지 못하는 오류가 있었다. 이는 전달하고자 하는 내용을 왜곡하여 오해를 일으키는 원인이 될 수 있다. 본 논문은 이러한 문제를 해결하기 위해 수화 아바타를 위한 수화 문장을 분석하고 수화 사전에 품사 정보를 넣는 방법을 제안한다.

본 논문의 2 절에서는 수화 문장 및 동음이의어 문제를 다루고, 3 절에서는 이를 해결하기 위한 수화 사전의 개선 방법을 제안한다. 4 절에서는 실험 및 프로그램 구현 결과를 보이고, 5 절에서 결론을 맺는다.

2. 수화 문장 및 동음이의어 문제

2.1 수화 문장에 대한 이해

일반 문장은 의사나 감정을 전달하기 위해 문자 언어로 나타낸 것이고 구와 어절, 단어, 형태소로 구성되어 있다[7]. 수화 문장은 청각장애인이 의사를 정확하게

전달하고 직관적으로 이해할 수 있도록 국어와는 다소 다른 독자적인 문법을 가지고 있으며, 수화 단어와 수화소, 지화로 구성되어 있다. 수화소는 수화 형태소라고도 하며 수화를 표현하는 최소 단위이다. 표 1은 일반 문장과 수화 문장의 예를 보여준다. 이와 같이 청각장애인이 수화를 할 때에는 조사와 시제, 접사를 무시하는 경우가 있다. 또한, 일반 문장에서는 어미 변화가 다양하지만 수화 문장에서는 그렇지 못하고 체언과 서술어의 위치가 달라질 수 있다[2].

표 1. 일반 문장과 수화 문장의 예

일반 문장	어머니께서 밥을 드신다 모레 눈이 올 예정이다
수화 문장	어머니 밥 모레 눈 아마

2.2 동음이의어에 대한 처리 오류

기존의 수화 시스템에서는 사용자가 입력한 문장을 어절로 분리한 후 형태소를 분석하고, 형태소가 수화 형태소이면 수화 단어 데이터베이스에서 해당 단어의 데이터를 탐색한다. 하지만, 동음이의어에 대해 의미가 다른 수화 단어 데이터를 탐색하여서 오류가 발생하는 경우가 있다. 이는 수화 아바타가 청각장애인에게 의미를 잘못 전달하고 오해를 일으킬 수 있다. 표 2는 동음이의어에 대한 처리 오류의 예를 보여준다. 기존의 수화 시스템에서는 “저(I)”의 수화 단어를 “저(That)”로 잘못 표현하는 경우가 있다.

표 2. 동음이의어에 대한 오류

문장 입력 1	제가 저 사람이 오겠다고 말하였다.
처리 오류	저(대명사) 저(관형사)
문장 입력 2	엄마는 자로 길이를 재고 아이는 자고 있다.
처리 오류	자(명사) 자(동사)

3. 수화 단어 사전의 개선

본 절에서는 동음이의어에 대한 처리 오류를 극복하기 위해 수화 단어 사전을 개선하는 방법을 제안한다. 이를 위해 우선 수화 단어 사전에 분석되는 형태소를 품사별로 분류하였다. 즉, 기존의 수화 단어 사전에 품사 정보와 수화 단어 코드, 형태소의 의미에 대한 설명을 넣었다. 그림 1은 동음이의어를 분류하고 수화 아바타의 동작을 구현하기 위해 수화 사전을 처리하는 단계를 설명한다.

그림 2는 개선된 수화 단어 사전의 형식을 보이고 그림 3은 수화 단어 사전의 예를 보인다. 동음이의어에

해당하는 “저(I)”와 “저(That)”는 품사별로 처리되고 값이 다른 수화 단어 코드로 구분된다. 구분된 수화 단어 코드에 따라 그림 4와 같이 수화 동작 데이터베이스에서 수화 동작 데이터를 검색하고 아바타의 동작으로 표현한다.

- 1) 문장 → 형태소 분석기
- 2) 수화 사전 내부 : (형태소, 품사, 형태소설명) → 수화 단어 코드
- 3) 수화 문장 재구성 : 수화 단어 코드 → 수화 단어
- 4) 수화 단어 동작 데이터 생성
- 5) 수화 아바타의 동작 구현

그림 1. 수화 사전의 처리 단계

형태소 품사 수화단어코드 형태소설명

그림 2. 수화 사전의 형식

까말	형용사	347	까말다, 까만
꺼말	형용사	347	
까지	조사	455	
까치	명사	456	
깨끗이	부사	457	깨끗하다, 깨끗해
자	명사	1038	
자	동사	1041	자다
자	어말어미	724	권유형
저	대명사	1337	나의 높임말
저	관형사	1336	지시 대명사

그림 3. 수화 사전에서의 품사 및 수화 단어 코드

347	검다 [347의 수화 동작 데이터]
455	까지 [455의 수화 동작 데이터]
456	까치 [456의 수화 동작 데이터]
457	깨끗하다 [457의 수화 동작 데이터]
724	자 [724의 수화 동작 데이터]
1038	자 [1038의 수화 동작 데이터]
1041	자 [1041의 수화 동작 데이터]
1336	저 [1336의 수화 동작 데이터]
1337	저 [1337의 수화 동작 데이터]

그림 4. 수화 동작 데이터베이스

그림 3과 4에서 “까말”과 “꺼말”은 같은 수화 단어 코드인 347로 처리되는데, 청각장애인은 건청인이 사용하는 “까말다”와 “꺼말다”, “흑색”, “검다”, “검은색”에 대하여 “검다”라는 수화 하나로 표현하기 때문이다[10].

4. 실험 및 프로그램 구현

동음이의어를 처리할 수 있는 수화 사전을 실험하고,

수화 단어 코드와 연관된 수화 동작 데이터를 사용하는 아바타 프로그램을 구현하였다.

4.1 MACH 한글 형태소 분석기

품사를 분류하기 위해 성신여대 심광섭 교수가 개발한 MACH 형태소 분석기[8,10]를 사용하여 동음이의어에 대한 처리를 실험하였다. 그림 5의 문장을 MACH 형태소 분석기에 입력하면 그림 6과 같이 형태소로 분석된다. 즉, 명사(NN), 동사(VV), 대명사(NP), 관형사(DT) 등의 품사가 분석되는데[9], “저(I)”는 “NN NP”로, “저(That)”는 “NN NP DT IJ”로 분석된다. 여기서 “저(That)”의 다음 품사가 조사가 아니면 “NN NP”의 처리를 생략하고, 수화 사전에서 “저”의 품사에 따라 코드가 다른 수화 동작 데이터를 찾는다.

저는 저 사람이 오겠다고 말하였다.
그림 5. 문장 입력의 예

저는 저 사람이 오겠다고 말하였다.
저는 :
 저/(NN NP) + 는/(JO)
 절/(VV) + 는/(EM)
저 :
 저/(NN NP DT IJ)
사람이 :
 사람/(NN) + 이/(JO)
오겠다고 :
 오/(VV VX) + 겠/(EP) + 다고/(EM)
말하였다. :
 말하/(W) + 었/(EP) + 다/(EM) + ./(SY)

그림 6. MACH 형태소 분석 결과[10]

4.2 수화 사전에 의한 수화 문장 재구성

① 수화 사전의 구성

입력한 문장을 수화 문장으로 변환하기 위해 수화 사전에 형태소와 품사 정보를 입력하였다. 수화 사전에 있는 약 1600 개의 형태소를 품사 별로 분류하면 표 3과 같다.

표 3. 수화 사전의 품사 개수

품사	개수	품사	개수
명사	1299	의존명사	5
동사	119	보조동사	4
형용사	81	조사	3
부사	39	감탄사	2
대명사	25	부호	2
어말어미	7	보조형용사	1
의존명사	6	수사	1
관형사	5		

② 수화 사전에서의 단어 중의성

2절에서 언급한 동음이의어에 대한 처리 오류의 원인을 규명하기 위하여 수화사전에서 중의성을 갖는 형태소를 분석하고 표 4에 정리하였다.

표 4. 중의성을 갖는 형태소 품사

품사	개수	품사	개수
명사	30	감탄사	2
동사	8	어말어미	1
대명사	4	조사	1
부사	2	의존명사	1
형용사	2	보조명사	1
관형사	2		

③ 수화 문장의 재구성

MACH 형태소 분석기에서 분석된 형태소는 수화 사전에 있는 형태소를 검색하여 수화 문장으로 재구성할 수화 형태소로 변환된다. 수화 사전에서 해당 형태소를 찾지 못하는 경우, 명사가 아니면 수화 형태소를 생략하고 명사의 경우에는 형태소를 초성과 중성, 중성으로 분리하여 지화 형태소로 바꾼다. 지화는 한글이나 영어를 알파벳으로 표시하는 손 제스처이다. 그림 5의 문장을 MACH 형태소 분석 및 수화 사전 검색을 거쳐 그림 7의 1)과 같은 형태소상 수화 문장으로 바꾸고, 수화 데이터베이스를 사용하여 그림 7의 2)와 같은 수화 단어 코드로 변환한 후, 그림 7의 3)과 같이 그래픽상 수화 문장을 표현하기 위한 수화 동작 데이터로 변환한다.

- 1) 형태소상 수화 문장: 저 저 사람 오 말하
- 2) 수화 단어 코드: 1337 1336 756 958 592
- 3) 그래픽상 수화 문장: 저 저 사람 오다 말

그림 7. 수화 문장의 재구성

4.3 프로그램 구현

그림 8과 그림 9에서 볼 수 있듯이 “저” 단어에 대한 수화 동작이 다르게 표현된다. 개발된 프로그램에서는 신체 요소의 동작 기호를 내장하여 연속적인 수화 동작과 목, 머리, 얼굴 요소 등 다양한 신체 요소의 동작을 함께 표현한다[11]. 또한, 수화 동작에서 손 모양이 잘 보여질 수 있도록 몸 전체의 회전이 자동으로 조절된다. 그림 8과 같이 “저(I)” 수화를 동작하면서 고개를 숙이고 자기를 낮추는 표현을 하고, 그림 9와 같이 “저(That)” 수화를 동작하면서 저쪽 방향을 손가락으로 가리키는 표현을 한다.

5. 결론

동음이의어에 대한 처리 오류로 인해 수화 아바타가

청각장애인에게 의미를 잘못 전달하고 오해를 일으키는 문제를 해결하기 위해 본 논문에서는 동음이의어를 품사별로 구분하여 처리할 수 있는 수화 사전의 개선 방법을 다루었다.

하지만, “사과”와 “배”, “말”, “밤” 등과 같이 품사도 같은 동음이의어에 대해서는 처리하지 못하며, 이는 앞으로 더 연구되어야 할 필요가 있다.

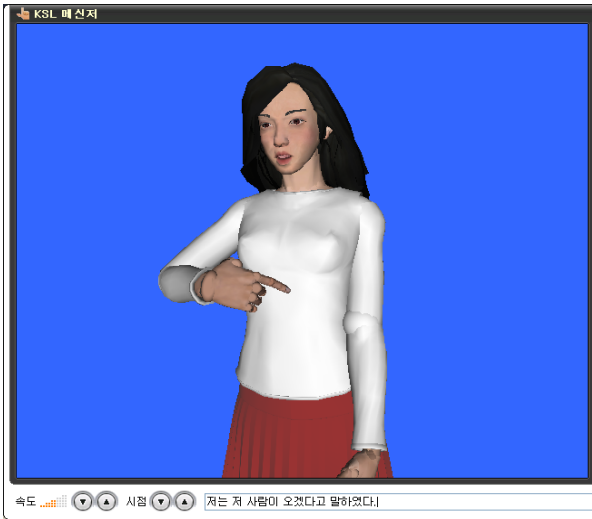


그림 8. “저(I)”의 수화 동작 구현



그림 9. “저(That)”의 수화 동작 구현

Acknowledgement

본 연구는 과학기술부/한국과학재단 우수연구센터육성사업의 지원으로 수행되었음 (R11-1999-008)

참고문헌

[1] “Sign language”, Encyclopedia, http://en.wikipedia.org/wiki/Sign_language
 [2] 장진석, 농아인과 사회, 한국수화에 대한 이해, 한

국농아인협회, 2000년

[3] Jong-Sung Kim, Won Jang and Zeungnam Bien, A Dynamic Gesture Recognition System for the Korean Sign Language (KSL), IEEE Trans. on Systems, Man, and Cybernetics vol. 26, no. 2, 354-359, 1996
 [4] 변증남, 장원, 김정배, 김대진, 조성식, 청각장애인 교육을 위한 수화 통역시스템의 개발, 한국과학기술원 보고서, 2000년
 [5] 김대진, 김정배, 장원, 변증남, TV 자막 신호를 이용한 한글 수화 발생시스템의 개발, 대한전자공학회 논문지, 39권, C1편, 5호, pp. 32-44, 2002년
 [6] 장효영, 오영준, 정성훈, 박광현, 변증남, 청각장애인을 위한 작업 현장용 한국 수화 메시지 보드의 개발, 한국 HCI 학회 2007 학술대회 논문집, pp. 393-398, 2007년
 [7] 서울대학교 사범대학 국어교육연구소, 고등학교 문법, 교육부, 1996년
 [8] MACH 형태소분석기, 성신여대 심광섭교수 홈페이지, <http://cs.sungshin.ac.kr/~shim>
 [9] 심광섭, 양재형, 인접 조건 검사에 의한 초고속 한국어 형태소 분석, 한국정보과학회 논문지 소프트웨어 및 응용, 31권, 1호, pp. 89-99, 2004년
 [10] 권경혁, 민홍기, 한글문장-수화 번역기를 위한 사전구성, 대한전자공학회 추계종합학술대회 논문집, 제 21권, 제 2호, pp. 841-844, 1998년
 [11] 오영준, 박광현, 장효영, 김대진, 정진우, 변증남, 하이퍼 수화문장을 사용한 수화 생성 시스템, 제 25회 한국정보처리학회 춘계학술대회, pp. 621-624, 2006년