

청각장애인을 위한 일기예보 수화방송 시스템 개발

*오주현, 전성규, 은준호 **김민호, 권혁철 ***김익태 ****김재현

*KBS 기술연구소, **부산대학교 컴퓨터공학과, ***프림포, ****서경대학교 산업공학과
*jhoh@kbs.co.kr

Development of Weather Forecast Sign Language Broadcasting System for the Hearing-Impaired

*Juhyun Oh, Seong-gyu Jeon, Junho Eun **Minho Kim, Hyuk-chul Kwon ***Iktae Kim

****Jaihyun Kim

*KBS, **Pusan National University, ***Primpo, ****Seokyeong University

요약

청각장애인을 위한 지상파방송 서비스 중 자막방송은 100%에 가까운 편성 비율을 달성하고 있지만, 화면을 가리는 수화방송은 5% 수준의 편성에 그치고 있다. 본 연구에서는 자막방송을 수화로 번역하여 그래픽 수화방송을 생성함으로써 수화방송의 비율을 높이고자 한다. 수화 단어들의 빈도를 파악하고 중요 단어부터 모션 캡처하기 위해 과거 3년간 일기예보 스크립트를 분석하였다. 자막방송 문장을 형태소별로 분석한 다음 중요 품사 위주로 단어 단위로 번역하고, 기 구축된 한국어 어휘의 미망을 이용하여 수화사전에 없는 유의어와 하위어를 대표어로 대체하였다. 기계번역 기술이 수화통역사의 수준을 따라잡을 수는 없지만 향후 수화방송도 선택적 서비스가 가능해지고 수화통역사의 수화방송이 모든 프로그램에 편성될 때까지 본 시스템이 보조적 시청 수단으로 사용 가능할 것이다.

1. 서론

공공재인 전파를 이용한 방송서비스는 누구라도 보편적으로 이용할 수 있어야 하며, 시청각장애를 가진 장애인도 예외가 아니므로 방송사는 장애인에게 방송접근권을 보장할 수 있도록 노력하여야 한다. 우리나라 장애인의 수는 2010년 기준으로 시각장애인 25만 명, 청각장애인 26만 명 및 난청노인 166만 명 등으로 해마다 크게 증가하고 있는 실정이다[1]. 2008년 장애인 관련 통계에 의하면 장애인의 문화 및 여가 활동 중 가장 많은 것은 'TV 및 비디오 시청'으로 나타났는데, 비율로도 전체의 36.1%를 차지해 두 번째로 높은 휴식(19.8%)보다 압도적으로 높았다[2].

이처럼 장애인 시청자의 생활에서 큰 비중을 차지하는 TV 시청을 위해 지상파 방송사는 자막방송, 수화방송, 화면해설방송 등의 장애인 방송 서비스를 제공하고 있다. 이 중 화면해설방송은 화면을 볼 수 없는 시각장애인을 위해 영상의 내용을 말로 설명하는 것이며, 자막방송과 수화방송은 청각장애인을 위해 음성 및 음향 정보를 각각 자막과 수화로 내보내는 것이다. 이 때 자막방송은 방송 영상에 영향을 주지 않도록 별도 텍스트 데이터(폐쇄자막, closed caption)로 전송되는 반면, 수화방송은 수화통역사 영상이 방송 제작 영상에 합성된 채로 송출된다(그림 1).

현재 KBS를 비롯한 지상파방송사의 자막방송 편성 비율은 거의 100%를 달성하고 있지만 수화방송 비율은 5% 수준에 머물고 있는데, 비용 문제도 있지만 폐쇄자막방송과 같은 선택적 서비스가 불가능한

수화방송의 특성이 가장 큰 원인이라고 할 수 있다. 자막방송은 프로그램 제작과는 전혀 무관하게 자막속기사가 입력하여 전송하고 원하는 사람만 수상기에서 선택하여 시청하면 되지만, 수화방송은 화면의 상당부분을 가리기 때문에 프로그램 기획 단계에서부터 고려하지 않



그림 1. 청각장애인 방송의 예. 자막방송(위)과 수화방송(아래).

으면 안 된다. 또한 수화방송을 포함하여 프로그램을 제작하더라도 장애인 시청자들에게는 수화 화면의 크기가 작고 위치를 조절할 수도 없기 때문에 불만이 제기되고 있는 실정이다.

표 1. 최근 지상파방송사 장애인방송 편성 비율 변화 (단위: %)

	2006	2007	2008	2009	2010	2011
자막방송	58	78	92.6	94.6	96	96.5
수화방송	3.6	2.9	3.8	4.0	5.1	5.2
화면해설 방송	4.7	4.9	5.3	5.5	6.0	6.2

표 1을 보면 수화방송은 편성 비율이 증가하긴 했지만 여전히 5% 정도로 미미한 수준인 데 반해, 자막방송은 2006년부터 크게 증가하여 거의 모든 방송시간대에 지원되고 있음을 알 수 있다[3]. 만약 자막방송으로부터 수화방송을 생성한다면 수화방송이 편성되지 않는 나머지 95%에 해당하는 방송시간대에도 수화방송을 제공할 수 있을 것이다.

수화방송을 자동 생성하는 기존 연구가 없었던 것은 아니다. 국내에서는 교통방송에서 교통정보를 수화로 전달하는 서비스를 제공한 바 있으며, 일본에서는 NHK에서 TVML(TV Making Language)를 이용하여 수화방송을 생성하는 연구가 진행 중이다[4].



그림 2. 국내외 캐릭터 수화 서비스. 위 왼쪽부터 교통방송, NHK, 중국의 캐릭터 수화방송[5]

2. 일기예보-수화 번역

이와 같은 수화방송을 생성할 때에는 자연스러운 그래픽 애니메이션과 더불어 정확한 한국어-수화 번역이 가장 중요하다. 그림 3은 전문 수화통역사에게 의뢰하여 일기예보 스크립트를 번역한 예를 보여준다[5].

뉴스	오늘은 전국적으로 하늘이 구름이 많은 상태입니다.
수화	오늘+전국+하늘+구름+많다(비수지기호).
뉴스	남부지방과 제주도에서는 이른 아침부터 비가 시작됐는데요.
수화	따뜻한+곳+도+제주도+곳+아침+부터+비+내리다.
뉴스	또 오늘 중부지방에서도 가끔 구름이 많은 가운데 비가 내리지 않는 지역 한낮기온은 서울 22도, 부산 23도 등으로 예년이맘때 기온보다 5, 6도 가량이나 웃돌 것으로 보입니다.
수화	또+오늘+전국+가운데+가끔+구름+많다+비(내리다)+않다+지역(곳,범위)+낮+온도+서울+22+부산+23+과거+때+온도+보다+5+6+온도(비수지기호)+아마(예측,-할것이다)
뉴스	현재 남서쪽 해상에서 계속해서 비구름이올라오면서 현재 전남 해안과 영남 일부지역에 약하게 비가 내리고 있습니다.
수화	지금+따뜻한+서쪽+바다+곳+계속+비+구름+오다+지금+전남+바다+도+강원도+따뜻한+곳+조금+비(내리다)
뉴스	주말마다 추위가 찾아오고 있습니다.
수화	일주일+마지막+춥다

그림 3. 일기예보 한국어-수화 번역 예

제안하는 캐릭터 수화방송 시스템에서는 자연어 처리(기계번역) 기술을 사용하여 이와 같은 수화 번역이 컴퓨터에 의해 자동으로 이루어질 수 있게 한다. 이를 위해 과거 3년간의 KBS 일기예보 스크립트를 분석하고 품사별로 단어 등장 빈도를 조사하였다.

표 2는 일기예보에 사용된 단어 중 명사, 동사, 형용사, 부사에 대해 가장 많이 등장한 어휘 10개씩을 보여준다. 표에서 알 수 있듯이 일기예보에 가장 많이 나타난 명사 단어는 '내일'인 것으로 조사되었다. '누적 빈도'는 단어를 중요도 순으로 정렬한 상태에서 해당 단어가지만 사전에 등록했을 때, 시스템이 전체 일기예보 스크립트의 몇 퍼센트를 처리할 수 있는지를 알려주기 때문에 중요한 숫자이다. 예를 들어 명사의 경우 발생빈도 10위인 '전국'까지 사전에 등록하고 모션 캡처 했다면 확률적으로 스크립트의 약 31%를 처리할 수 있다. 부사의 경우 상위 10개만 등록하더라도 전체 스크립트의 69%를 처리 가능하다고 볼 수 있다.

현재의 캐릭터 수화방송 시스템은 일기예보에 적용할 수 있게 어휘와 캐릭터 모션 등을 추가하여 개발 중이다. 번역 시스템은 한국어 텍스트가 입력되면 이를 형태소 단위로 분석하여 중요하지 않은 조사 등을 제거하고, 동의어와 하위어를 상위어(대표어)로 대체하는 방식으로 수화로 변환한다. 이를 위해 그림 4의 한국어 어휘의미망(KorLex)[6]이 활용된다.

집 1, 주택 1, 주거 1, 가옥 1

house 1 (a dwelling that serves as living quarters for one or more families: "he has a hot get out of the house")

KORLEX	WORDNET	KORLEX
상위개념	집 1, 가옥 1	하위개념
	집 1, 주거 1	비치하우스 1
	집 1, 건물 1	향숙집 1
	집 1, 건물 1	방갈로 1, 단층집 1, 단층 2
	가공품 1, 가공품 1, 인공품 1	오두막집 1, 오막살이 1
	건조물 1, 구조물 1, 구조 2	상권 1
	방갈로 1, 건조물 1, 건축물 1, 건물 1	남학생사교클럽회원 1
	집 1, 주택 1, 주거 1, 가옥 1	시골저택 1
	주거시설 1	독립주택 1, 단독주택 1, 주택 2
	집 1, 거처 1	물하우스 1
	집 1, 주택 1, 주거 1, 가옥 1	두세대용_주택 1

그림 4. KorLex 예

3. 수화 캐릭터 애니메이션

캐릭터 애니메이션을 위해서는 그림 5와 같이 수화 모션을 캡처해

표 2. 일기예보 단어 빈도 (품사별 상위 10개)

명사			형용사			동사			부사		
어휘	발생(회)	누적빈도(%)	어휘	발생(회)	누적빈도(%)	어휘	발생(회)	누적빈도(%)	어휘	발생(회)	누적빈도(%)
내일	3,295	5	높다	1,277	14	기다	1,582	11	조금	742	11
낮	2,664	9	많다	2,450	27	내리다	1,451	21	점차	702	21
오늘	2,602	13	있다	3,575	40	일다	854	26	대체로	628	30
기온	2,598	16	따다	4,416	49	오다	844	32	또	576	38
비	2,092	20	낮다	5,214	58	오르다	821	38	다소	470	45
서울	1,747	22	강하다	5,823	65	이어지다	737	43	더	417	51
아침	1,611	24	비슷하다	6,303	70	끼다	712	47	가끔	284	55
지역	1,585	27	크다	6,679	74	받다	628	52	다시	269	59
중부지방	1,426	29	길다	7,022	78	지나다	493	55	특히	204	61
전국	1,293	31	춥다	7,335	81	떨어지다	440	58	매우	164	69

서 데이터베이스에 가지고 있어야 한다. (주)프림포에서 기존 보유하고 있던 약 2,700개의 수화 모션에 일기예보를 위해 약 500개의 수화를 추가함으로써 전체 약 3,200개의 수화 모션을 보유하게 되었다.

30 프레임을 기준으로 15 프레임을 모션 블렌딩 구간으로 정의한다. 초록색 선이 노란색과 파란색 사이에 블렌딩된 모션을 나타낸다.

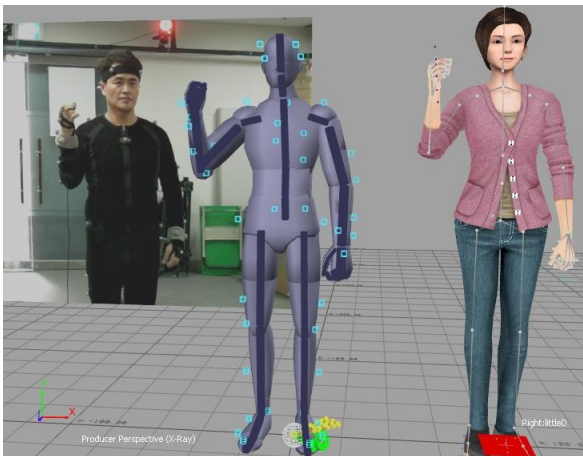


그림 5. 수화 모션 캡처 화면. 왼쪽부터 연기자, 뼈대(skeleton), 최종 캐릭터

캐릭터 애니메이션에 있어서는 두 개의 모션이 서로 다른 관절 운동을 표현하고 있을 때 이를 부분적으로 추출하고 합성하여 새로운 모션을 갖는 데이터를 생성할 필요가 있다. 예를 들어 수화 단어 끝 부분과 다음 단어의 시작 부분을 보다 부드럽게 연결을 시켜 마치 하나의 문장으로 표현 될 수 있도록 구현할 수 있는데, 식 (1)은 파라미터 u (blending factor)에 따라 두 곡선 사이에 블렌딩(blending)된 곡선을 생성하는 것으로, 그림 6은 두 스플라인(spline) 곡선 사이의 기본 블렌딩을 보여 주고 있다. 두 개의 사용자 모션을 나타내는 곡선에 이것을 적용시킬 때, 두 모션에서 보간(interpolation)된 새로운 모션을 생성할 수 있다.

$$d^*j = ub^*j + (1-u)c^*j, \quad 0 \leq u \leq 1 \quad (1)$$

그림 7은 모션 블렌딩을 통한 중간 모션을 생성한 것이며, 초당

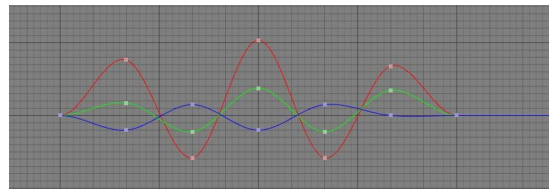


그림 6. 모션 블렌딩 그래프

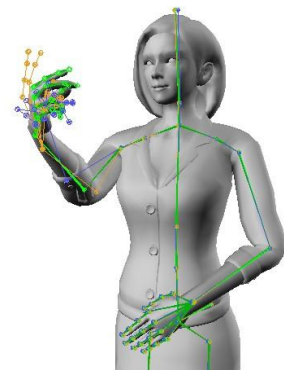


그림 7. 스켈리톤 모션 블렌딩

4. 실험 결과

그림 8은 일기예보 수화방송 실험 소프트웨어가 작동하는 모습을 보여준다. 일기예보 영상을 재생하면 해당 기사의 스크립트를 실시간 번역하여 캐릭터의 애니메이션으로 재생한다. 수화 애니메이션 화면을 보면 “서울의 아침 기온이”와 같은 한국어 일기예보 스크립트가 “서울, 아침, 기온”과 같이 수화의 해당 단어로 번역된 결과를 볼 수 있다. 번역모듈에서 번역에 실패하거나 수화모션 데이터베이스에 해당 단어의 애니메이션이 없는 경우에는 ‘N’으로 표시하도록 하였다.

표 3은 내부 말뭉치(corpus)와 외부 말뭉치에 대해, KorLex[7]를 적용했을 때와 그렇지 않은 경우의 결과를 보여 준다. 표에서 내부 말뭉치는 수화 기본어 사전 구축을 위해 분석한 말뭉치를 뜻하며, 외

부 말씀치는 수화 기본어 사전 구축 후 평가를 위하여 추가로 확보한 말씀치이다. KorLex를 적용하여 번역률이 높아진 결과를 확인할 수 있으며, 구체적인 예로는 수화 기본어 사전에 없는 단어인 ‘토사’가 상위어인 ‘흙’으로 번역되는 경우 등을 들 수 있다. KorLex를 적용하더라도 번역에 실패하는 예로는 ‘즐라왓’과 같은 태풍의 이름이라든지, 사전에 등록되지 않은 저빈도 단어가 등장했을 경우 등이 있다.



그림 8. 일기예보 수화방송 실험 소프트웨어. (왼쪽)수화 애니메이션, (오른쪽 위)번역 화면, (오른쪽 아래)원본 일기예보 영상

표 3. 한국어 어휘의미망(KorLex) 적용 전후 결과 비교

말뭉치 구분	말뭉치 크기 (어절)	번역률(%)	
		KorLex 적용 전	KorLex 적용 후
내부 말뭉치	82,303	95.87	96.17
외부 말뭉치	1,448	88.60	90.68

5. 결론

한국어로 작성된 일기예보 방송 스크립트를 수화로 번역하고 캐릭터 애니메이션으로 표현하여 수화방송을 생성하는 시스템을 제안하였다. 이를 위해 과거 3년간 방송된 일기예보 스크립트를 분석하여 활용하였다. 자막방송 문장을 형태소별로 분석한 다음 중요 품사 위주로 단어 단위로 번역하고, 한국어 어휘의미망을 이용하여 수화사전에 없는 단어를 대표어로 대체하였다. 기계번역 기술이 수화통역사의 수준을 따라잡을 수는 없지만 향후 수화방송도 선택적 서비스가 가능해지고 수화통역사의 수화방송이 모든 프로그램에 편성될 때까지 본 시스템이 보조적 시청 수단으로 사용 가능할 것이다.

지금은 미리 저장된 일기예보 영상과 그에 동기화된 스크립트를 사용하였지만 향후 모든 프로그램에 사용 가능하도록 번역 소프트웨어와 어휘, 그리고 수화 모션 등을 확장할 계획이다.

이와 같은 확장이 이루어지면 본 연구의 최종 목표인 DTV 수화방송 서비스 이전에 그림 10과 같이 스마트 DMB[7]의 한 서비스로 적용하는 것을 검토 중이다. DMB는 DTV보다 영상의 지연(delay)이 큰데, 광고롭게도 이 지연시간이 자막 입력에 의한 3초가량의 폐쇄자막의 지연시간과 비슷하게 맞아 떨어지기 때문에 스마트 DMB에서는

어느 정도 동기화된 수화방송 서비스가 가능할 것으로 생각된다.

앞으로 수화번역의 정확도와 어휘 및 모션 확장 등 앞으로 풀어야 할 문제가 많지만 일기예보 수화 생성을 통해 그 가능성을 확인하였고, 많은 장애인 시청자들에게 조금이나마 도움이 될 수 있도록 연구개발을 계속해나갈 계획이다.



그림 10. 스마트 DMB 서비스 활용 예상 그림

참고 문헌

- [1] 홍종배, "장애인 방송접근성 표준화 동향," TTA Journal, 2011년.
- [2] 한국보건사회연구원, "2008 장애인실태조사," 2008년.
- [3] 방송통신위원회, "장애인방송 가이드라인(안)," 장애인방송 공청회, 2011년 5월.
- [4] H. Kaneko 외, "Sign language animation using TVML", VRC AI 2010.
- [5] <http://blog.daum.net/starbiz/12731257>
- [6] (주)프림포, "수화 캐릭터 애니메이션 연구," 미출간, 2012년 12월.
- [7] 윤애선 외, "한국어 어휘의미망 「KorLex 1.5」의 구축," 정보과학회논문지, 2009년.
- [8] <http://www.kbs.co.kr/dmb/clover/smartdmb.html>