

ETRI의 대화체 음성언어번역 연구방향

양재우, 이영직, 한민수
 한국전자통신연구소 휴먼인터페이스연구부
 305-600 대전지하철 유성구 유성우체국 사서함 106호
 jwyang@media.etri.re.kr

Research Trends on Spontaneous Speech Translation at ETRI

Jae-Woo Yang, Youngjik Lee, and Min-Soo Han
 Human Interface Technology Department, ETRI P.O. Box 106, Yusung, Taejeon, 305-600, Korea

요약

본 발표에서는 한국전자통신연구소의 대화체 음성언어번역 연구방향을 기술한다. 한국전자통신연구소에서는 1979년 공중전화 관련 연구를 필두로 음성분야의 연구를 시작하여 그간 음성인식, 자동통역 등을 연구하였다. 특히 1991년부터 KT와 공동연구체계를 수립하여 특정 분야를 대상으로 음성언어번역 연구를 수행하는 중이다. 이러한 경향을 바탕으로 한국전자통신연구소에서는 국내외 공동연구를 통해 대화체 음성언어 번역 연구를 수행하고자 한다. 앞으로 다중매체통신이 더욱 보편화 될 것으로 다중매체 환경을 고려하여 연구를 수행할 계획이다. 아울러 이러한 연구의 부산물을 상용화 하는데 노력을 경주할 계획이다.

1 서론

한국전자통신연구소는 1983년부터 음성분야를 연구하여, 그간 문서낭독기, 전자비서시스템등의 연구실적을 이룩하였다 [1, 2, 3]. 또한 국내외 연구기관과 공동연구를 추진 하였으며, 특히 KT와 공동연구 체계를 1991년부터 수립하여 정기적인 기술간담회를 통해 기술을 교류하고 있다.

본 논문에서는 한국전자통신연구소의 음성분야 연구동향을 기술한다. 그 큰 흐름을 살펴보면 대화체 음성언어 번역, 국내외 공동연구, 음성처리기술의 상용화로 볼 수 있는데, 이를 각 절 별로 살펴보고, 마지막에 결론을 맺는다.

2 대화체 음성언어 번역

본 연구의 최종 목표는 고속 정보통신망에서 지원되는 다중매체 환경에서 자국어를 사용하는 대화체 음성언어 통신 기술의 개발이다.

기존의 통신망에서는 공중전화망을 통해 사람 간에 의사를 전달하였다. 그런데 이 경우 오직 음성만을 사용하게 되므로, 약도나 물건 외 모양 등을 전달하려면 어려운 점이 많이 있었다. 아울러 이 때 사

용되는 음성언어의 형태도 자연스럽게 대화체와는 다르게 된다. 그런데 다중매체 환경에서는 동화상, 정치화상, text등도 전달되므로 의사전달의 통로가 다양해진다. 즉 상대방의 몸짓이나 얼굴 표정을 통해 상대방의 의사가 전달되며, 아울러 약도나 물건의 모양이 화상으로 전달되므로, 이에 따른 음성언어의 형태가 매우 자연스러워 질 것이다. 이러한 측면에서 볼 때 대화체 음성언어의 연구가 매우 필요하다 하겠다.

음성을 이용한 의사 전달은 사람이 가진 의사 표현 수단중 가장 자연스러운 것이라 할 수 있다. 음성을 이용하여 통신을 할 경우 아플 이용하기 위해 특별히 훈련을 할 필요가 없다. 아울러 정보 전달이 매우 빠르다는 장점도 가지고 있다. 여기에 다중매체 통신이 부가되면 매우 자연스러운 통신환경이 될 것이다. 따라서 다중매체 환경에서의 대화체 음성언어 번역 통신 기술개발은 명실공히 차세대의 음성응용 서비스가 될 수 있다.

대화체 음성언어 번역 연구는 선진 각국에서 많이 수행되고 있다. 예를 들어 미국 Carnegie Mellon 대학의 JANUS 시스템은 영/일, 영/독 음성언어 번역이 가능하다. 일본의 ATR에서 개발된 ASURA 시스템을 일/영 번역이 가능하고, 유럽의 SRI Cambridge에서는 영/스웨덴 음성언어번역 시스템을 개발하였다. 이러한 음성언어 번역 연구는 국제 공동 연구의 형태를 주로 취하는데 일본의 ATR, 미국의 CMU, 독일의 Karlsruhe 대학이 참여한 C-STAR, 이것이 확장된 C-STAR II 등이 좋은 예이다.

대화체 음성언어 번역 기술은 크게 대화체 음성인식 대화체 언어 이해, 대화체 언어번역, 대화체 음성합성으로 나눌 수 있다. 대화체 음성인식은 대화체 음성에서 자주 나타나는 속도, 억양등의 변화를 처리해야 하며, 불확실한 발성을 인식할 수 있어야 한다. 대화체 언어 이해는 발성 중간에 삽입되는 무의미 단어를 처리할 수 있어야 하며, 대화문장에 자주 나타나는 생략, 도차, 대응어 등을 이해하여 그 의

ETRI의 대화체 음성언어번역 연구방향

미를 추출해야 한다. 대화체 음성합성은 운율을 살려 자연스러운 음성을 합성해야 한다.

3 국내의 공동연구

한국전자통신 연구소에서는 이러한 연구를 효율적으로 추진하기 위해 국내의 공동연구를 체계화 하고자 한다. 즉 국내 공동연구를 위해 음성언어 번역 연구센터를 운영함으로써 음성관련 연구의 기술 축적을 도모한다. 국제 공동연구는 1994년 6월에 가입한 C-STAR 활동, 1993년 12월에 양해 각서를 교환한 일본 ATR과의 기술교류, 미국 CMU와의 공동개발등을 활용하여 기술 경쟁력을 재고할 계획이다.

음성언어 번역 연구센터는 국내 연구인력의 결집을 도모하고, 음성관련 연구의 기술 축적 및 재발용을 위해 1995년에 설립할 예정이다. 이 센터를 통해 기반기술 및 음성 DD 등 관련 기반 데이터를 국가적으로 공유하고자 한다. 이 연구센터에는 시스템 통합, 음성인식, 음성번역, 음성합성 그룹을 두며 총 30-40명의 규모로 운영할 계획이다.

국제공동연구는 크게 C-STAR 활동, 일본 ATR과의 공동연구, 미국 CMU와의 공동연구로 나누어 볼 수 있다. ETRI가 1994년에 가입한 C-STAR-II는 Consortium for Speech Translation Advanced Research의 약자로서 대화체 음성언어 번역 연구를 위한 국제간 컨소시움이다. 일본의 ATR, 미국의 CMU를 위사하여 MIT, 프랑스의 LIMSI, 이태리의 IRST, 영국의 SRI Cambridge, 독일의 SIEMENS, DFKI, Karlsruhe 대학이 회원이며 1996년 및 1999년에 공동 실험을 할 예정이다. C-STAR는 년 1회 회의를 통해 기술을 교류한다. 일본 ATR과는 한일간 음성언어번역 기술을 공동으로 연구하며, 9 개월마다 개최하는 기술간담회 및 연구원의 교환 파견을 통해 기술교류를 다질 예정이다. 미국의 CMU와는 한/영 음성언어 번역을 연구하며 software 공동이용 및 연구원 파견을 통해 기술을 교류한다.

이러한 연구를 바탕으로 한국전자통신 연구소는 1996년에 대화체 음성언어 번역의 요소기술을 시범하고, 1999년에 국제간 음성언어번역을 시범할 계획이다.

4 음성처리 기술의 상용화

음성언어 번역 연구과정에서 생기는 실용화가 가능한 연구 부산물로는 음성 합성 기술, 음성 입력 기술, 음성 dialing 기술 등을 들 수 있다. 기술 개발의 궁극적 목적이 실용화에 있으므로 이러한 연구 결과를 실용화하는 노력은 매우 중요하다. 이러한 기술을 상용화하면 음성 합성기, 음성 입력기, 음성 동작 기기, 음성 dialing 서비스 등을 제공할 수 있다. 이러한 기술은 국제간 자중예약 등 교류가 가

치 통신 서비스를 가능케 할 뿐만 아니라, 국제간 영상회의, 장애인 통신 등에 활용될 수 있다.

5 결론

본 논문에서는 한국전자통신연구소의 음성분야 연구동향을 기술하였다. 즉 다중매체환경에서의 대화체 음성언어번역 통신기술개발을 국내의 공동연구를 통해 추진하고, 실용화 가능한 연구결과를 최대한 살려 음성분야 기술의 국제 경쟁력을 높일 예정이다.

감사의 글

저자들은 본 연구와 관련한 한국통신의 꾸준한 지원에 감사로 드립니다. 아울러 저자들은 한국전자통신연구소의 음성분야 연구방향 설정에 중요한 역할을 해 주신 이원용 부소장님과 강철희 박사님께 큰 감사의 뜻을 표합니다.

참고 문헌

- [1] 이용주, "자동통역전화의 기술 현황 및 과제," 전자공학회지, 1993.7.
- [2] Sang-Hun Kim, Minje Zhi, Un-Cheon Choi, and Hee-Il Han, "Application of TD-PSOLA technique T-t-S conversion," *Proc. IWSP'98*, 1993.9.
- [3] Young Mok Ahn, Hoi Rin Kim, and Kyu Woong Hwang, "Development of a simple VQ-based pre-processor," to appear in *Proc. ICSPAT*, 1994.10.