

자동통역전화의 기술현황 및 과제

李 勇 柱

韓國電子通信研究所 自動通譯研究室

I. 서론

음성언어의 공학적 연구에서는 주로 음성언어의 특징을 분석하고(음성분석), 이러한 특징을 이용하여 기계가 음성을 알아듣고(음성인식), 그 요구사항을 처리하여 다시 음성으로 알려주는(음성합성)기술이 그 주된 연구분야이다. 요구사항의 처리내용에 따라 여러 가지 응용이 있을 수 있는데 그 처리내용이 언어간의 기계번역일 경우에는 원리적으로 음성언어의 자동통역이 가능하다. 자동통역전화 시스템은 언어가 서로 다른 사람들 간의 대화내용을 쌍방의 언어로 동시통역 해주는 통역사로 생각할 수 있으며 통화가자 한 말을 인식하는 음성인식, 인식된 말을 상대방의 언어로 번역하는 기계번역, 번역된 문자를 음성으로 출력해 주는 음성합성의 3가지 핵심기능을 양방향으로 실시간에 처리하는 그림 1과 같은 시스템이다. 이러한 시스템은 언어의 장벽을 극복해주는 지능적인 통신 서비스를 가능하게 하는 것으로, 음성언어정보 처리연구의 궁극적인 목표의 하나라 할 수 있다.

및 연구해야할 과제에 대해서 소개하고자 한다.

II. 자동통역전화시스템의 구성요소^[1]

자동통역전화시스템은 기본적으로는 발생된 회화음성의 내용을 번역하는 음성번역(speech translation)장치를 서로 다른 언어사이에 쌍방향으로 설치하여 구성한다. 음성번역시스템의 일반적인 구성은 그림2와 같다.

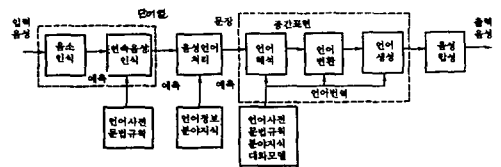


그림 2. 음성번역의 일반적 구성^[1]

말하는 사람이 연속해서 발생한 음성은 먼저 음소 단위로 인식하고, 이 음소들의 연결로써 연속음성을 인식하여 문자열로 변환한다. 이때 사전과 문법규칙이 이용된다. 음성인식의 결과는 몇개의 후보로 나타내며 각각 신뢰도(또는 유사도) 값을 가지고 있다. 이렇게 얻어진 결과로부터 구문을 주로한 언어정보와 해당분야에 관한 지식을 이용하여 알맞은 문장이 선택된다. 문장은 구문, 의미 또는 단어의 용법과 분야에 관한 지식등에 의해 언어해석되어 중간표현이 얻어진다. 이때 대화모델을 이용해서 생략된 것들에 대한 보완도 이루어진다. 얻어진 중간표현은 상대국 언

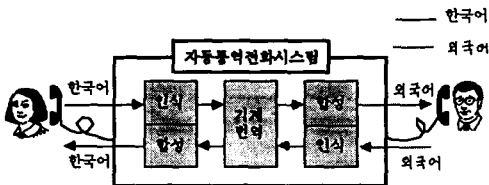


그림 1. 자동통역전화시스템의 개념도

본고에서는 국내외의 자동통역전화 관련 연구 현황과 함께 시스템 구성을 위한 요소 기술별 기술수준

어의 중간표현으로 변환되어 상대국 언어를 생성하고 이것을 음성으로 바꾸어 출력한다. 여기서 특히 중요한 것은 음성인식과 언어번역간의 음성언어통합처리이다. 인간이 한 말은 언어의 용법적인 제약 아래에서 발생된다는 것을 전제로 하여 음성통합처리에서는 음성인식부에 대해서 어떤 단어가 쓰이는가를 톱다운으로 예측하거나, 음성인식 결과로 나온 복수의 후보 중에서 적절한 문장을 선택한다. 분야예측에서는 예를 들어 호텔예약 등과 같이 특정분야를 대상으로 할 경우, 예약에 필요한 날짜, 이름, 화제나 용어 등의 데이터를 대상으로 하며, 이를 데이터베이스화 해두고 언어번역 및 음성언어처리에 이용한다.

한편 소규모의 음성번역시스템을 구성할 경우에는 음성인식부와 언어해석부가 하나로 묶여지는 경우가 보통이다. 이것은 연속음성을 언어정보와 분야지식을 이용하여 이해하고 직접 중간표현을 도출하는 것이다. 언어 및 분야의 모든 지식원을 통일적이고 집중적으로 이용하므로, 언어표현이나 어휘가 한정된 경우에 효과적이다. 구성은 그림 3와 같다. 이것을 일반적인 시스템과 비교할 때, 문장레벨의 음성인식과 언어해석을 통합하여, 발생된 문장음성을 인식한다. 그 결과로, 의미적 중간표현을 직접 만들어 내는 것이 다르다. 소규모 시스템에서는 특정분야 만으로 한정된 음성인식에 적합한 언어처리를 하기 때문에 언어표현이 제한적이다.

Ⅲ. 외국의 동향

음성 및 언어정보처리의 종합기술이라 할 수 있는 자동통역전화에 관한 연구는 미국, 유럽, 일본 등에서 음성 인식, 음성합성, 기계번역 등 요소 기술 별로 각기 활발하게 연구되어 왔고 최근에는 최종목표를 자동통역전화의 구현으로 정하고 일부에서는 제한적인 음성번역 또는 자동통역의 실험시스템들을 발표하였다.

음성번역을 목표로 한 최초의 실험 시스템은 영국의 British Telecom 연구소에서 1987년에 발표한 SALT(Speech Automatic Language Translation)으로 음성인식과 음성합성, 그리고 정형문 변환에 의해 호텔예약 등 1000단어정도로 구성된 400여 종류의 상용회화문을 대상으로 하였다.^[2] 이 시스템의 프로토타입이 제네바에서 열린 TELECOM '87에서 시연되었다. 이 시스템 하드웨어의 간단한 블록도는 그림4와 같다. 시스템간의 데이터는 데이터 모뎀에 의해 PSTN으로 연결된다. 대상문장을 단어별로 끊어서 확실하게 발생하고 끝에 "ENTER" 라고 발생하면 고립단어 인식시스템에 의해 이들 단어를 인식하여 발생된 단어들로 구성된 문장을 선택한 후 그 결과를 합성음성 및 화면상의 표시로 동시에 나타내 준다. 이때 발생자가 "yes", "no"로 결과를 확인하며 "no" 일 경우는 다시 처음부터 발생하고 "yes" 일 경우는 인식된 정보를 수신 측으로 전송한다. 그러면 수신 측에 상대국어에 해당하는 문장으로 변환 후, 합성음성 및 화면상의 표시로 변환된 문장을 출력한다. 이 시스템은 문장단위로 연속 발생된 것이 아니며 번역도 문장 대 문장으로 변환하는 초보적인 것이나 음성번역의 개념적인 형태를 실시간에 보여주었다는데 의미가 있다.

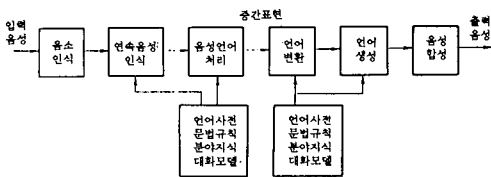


그림 3. 소규모 음성번역의 구성^[1]

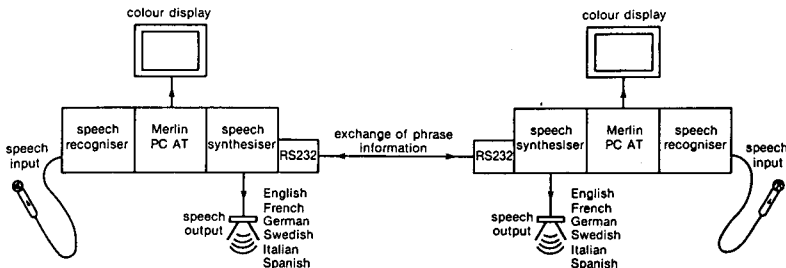


그림 4. SALT (Speech Automatic Language Translation) 시스템의 구성^[2]

1960년대부터 음성언어 정보처리의 응용기술연구를 꾸준히 해 왔던 일본은 1984년 우정성이 자동통역전화시스템 개발을 위한 국제공동연구를 제안하였고^[3], 1986년에 자동통역전화 만을 전문적으로 연구하는 자동번역전화연구소를 ATR(Advanced Telecommunications Research) 산하에 설립하였다.^[4] ATR에서는 1986년부터 15년간 900억엔의 예산규모로 국제회의가 가능한 일영 자동번역전화시스템 개발 계획을 마련하였고^[5] 그 첫단계 7년의 연구를 1993년 3월에 마치고 4월부터는 신설된 음성번역통신연구소 (Interpreting Telecommunications Research Laboratories)로 이어져 연구가 계속되고 있다.^[6]

1989년에 연구의 중간결과로 구성된 시범시스템이 "SL-TRANS"로서 이 시스템은 약 400단어 정도로 구성된 "국제회의 참가신청"에 관한 회화를 테스트로 하며, 화자적응기능을 이용한 HMM-LR 문절연속음성인식(문절은 우리말의 띄어쓰기 단위인 어절과 유사한 단위임.), 의도전달방식의 대화문 일영번역 및 영어음성합성으로 구성되었다. 그림 5에 그 구성을 보였다.

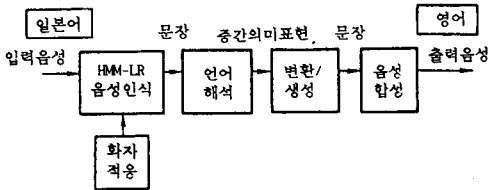


그림 5. 일영 음성언어 번역 실험시스템 (SL-TRANS)의 구성^[1]

미국의 CMU에서는 일본 ATR의 연구와 연계하여 영어를 독일어와 일본어로 번역해주는 "JANUS"라는 실험시스템을 1991년에 발표하였다. 이 시스템은 ATR과 같이 국제회의의 등록 및 문의를 테스트로 하고 있으며, 통계적 방법인 LR parser를 이용한 JANUS1 과 신경망에 의한 connectionist parser를 이용한 JANUS2의 두 방식으로 나누어 구성하였다. 그 결과, connectionist parser가 부정확한 음성인식 결과로도 LR parser에 비해 더 정확한 번역 결과를 나타내었다.^[7]

또한 이들과 연계하여 독일의 Siemens AG와 Karlsruhe대학교의 연구팀들도 독일어를 중심으로 영어와 일어로 번역하는 시스템을 구성하였고, 올해

들어 이들 미국, 독일, 일본의 3국간에 통합시범을 보인바 있다.^[8]

통합시범시에 ATR에서 새로 구성된 시스템으로 ASURA(Advanced Speech Understanding and Rendering System of ATR)가 있다.^[8] 이 시스템도 SL-TRANS 와 같이 국제회의 문의를 대상으로 하며 음성인식부와 언어번역부 모두 1500단어정도의 어휘를 다루고 있다. 먼저 음성인식부에서는 문절간에 포즈를 두어 문장단위로 발생하며 문절음성인식에는 SSS-LR 방식을 이용하였고 음소모델은 축차상태분할(SSS)알고리즘에 의해 자동 생성된 Hidden Markov Net(HMnet)를 사용하였다. 문절단위로 발생된 문장음성을 인식하기 위해서 음소환경에 의존하는 일반화 LR구문해석법에 의해 음소모델을 구동하고 있다. 문절내의 문법과 문절간의 문법을 사용하여 구문적으로 타당한 문장후보를 음성인식스코어 순으로 출력한다. 실시간 HMnet출력의 유사도 계산은 DSP로 구성된 고속프로세서로, 구문해석은 워크스테이션(HP9000/750)상에서 수행한다. 언어번역은 해석, 변환, 생성의 3부분으로 되어있다. 언어해석부에서는 일본어의 회화언어에서 자주 나오는 생략표현 및 간접적인 의뢰와 같은 다양한 표현을 다루기 위해 구문, 의미, 화용론적 제약을 Unification Grammar로 기술하고 있다. 변환부에서는 처리계와 변환규칙을 완전히 분리해서 다른 언어의 번역시스템 개발도 쉽도록 되어있다. 생성부에서는 Unification Grammar를 기본으로 하여 선언적인 규칙을 기술하도록 되어있고, 관용표현에도 유연하게 대응할 수 있는 구구조(phrase structure) 규칙을 채용하고 있다. 이러한 처리들은 워크스테이션(HP9000/750)상에서 수행한다. 일본어 음성입력에 의한 영어번역율은 90%정도를 얻고 있다.

한편, 일본의 NEC는 1991년에 영어와 일본어간의 쌍방향 자동번역시스템을 발표하였다. 이 시스템은 연속음성인식, 간단한 기계번역, 그리고 음성합성기술을 결합한 쌍방향 음성통역시스템으로, 연주회의 표 예약, 관광안내와 같은 정형적인 회화를 대상으로 하고 있다.

또, ATT Bell 연구소에서는 1991년에 연속음성인식, 언어처리 및 음성합성을 결합한 쌍방향 음성번역 실험을 영어와 스페인어 간에 은행에서의 간단한 회화를 대상으로 수행한 바 있다.

이밖에도 독일의 Ministry of Research and

Technology(BMFT)에서는 대화시에 즉석에서 사용할 수 있는 휴대용 통역지원시스템을 2000년까지 개발하기 위한 'Verbmobil'이라는 프로젝트를 준비하고 있다.

여기서는 2가지의 시스템을 목표로 하고 있는데, 첫번째는 사용언어가 서로 다른 두 사람이 대화할 때 각기 상대국어로 통역해주는 것이며, 또 하나는 두 사람 모두 모국어가 아닌 공통 언어(공용어)로 대화할때, 모르는 단어나 구 또는 문장을 부분적으로 통역해주는 시스템이다. 대상 텍스트의 후보로 "계약협상" 또는 "설계협상" 등이 검토되고 있다.

IV. 국내의 현황

국내에서도 지능적이고 편리한 통신서비스의 하나로써 언어의 장벽을 극복하기 위한 자동통역전화의 15년에 걸친 장기개발계획이 한국통신의 사업계획으로 마련되었다. 이는 국내의 음성언어 정보처리연구의 일대 전기를 마련해 주는 획기적인 일로서, 우선 한일간의 자동통역전화를 1차 목표로 하여 1991년부터 7년간 이에 필요한 요소기술을 한국전자통신연구소와 한국통신의 소프트웨어 연구소가 함께 개발하고 있다. 이 연구는 2005년 자동통역전화 프로토타입 실현을 위한 1단계 연구이다. 1단계 7년(1991 ~ 1997)의 요소기술 개발에 이어 2단계 (1998 ~ 2000)에는 자동통역전화의 실험모델을, 3단계(2001 ~ 2005)에는 한정된 텍스트의 자동통역전화 시스템을 개발하는 것을 목표로 하고 있다.^[9]

현재 세부개발계획의 수립에 이어 각 분야별 요소 기술 연구가 진행되고 있다.^[10] 음성인식 및 번역은 호텔 문의 및 예약을 텍스트로 한 문장음성인식 및 번역 시스템 구현을 목표로 하여 대어휘 인식을 위한 음소인식 기술, 문장음성인식을 위한 언어처리기술, 기계번역기술들이 개발되고 있다. 초기단계에는 200여 단어로 구성된 한정된 문법하의 문장음성 인식 및 번역을 목표로 한다.

음성합성의 경우도 자연스런 음질과 다양한 음색표현을 목표로 한 합성방식 및 규칙에 관한 연구가 진행되어 700여 개의 반음절을 LSP(Line Spectrum Pair)파라미터로 표현하는 방법을 이용하여 어휘에 제한이 없는 한국어 합성시스템을 실현한 바 있고,

명료성을 더욱 개선 시키기 위해 약 2000여개의 CDU(Context Dependent Unit)와 파형 단위의 PSOLA (Pitch Synchronous Overlap and Add) 방식에 의한 프로토타입이 구성되어 높은 명료성을 보였다. 이 시스템은 운율처리를 위한 언어지식의 효율적 활용, 완성도 높은 글자 읽기 규칙의 적용 등으로 여러 가지 가능성을 보였다. 현재, 남녀 음성의 합성 및 자연성 개선을 위한 연구가 계속되고 있으며 더욱 고품질의 합성이 가능한 방법들도 연구되고 있다. 또한 음성에 의한 질의 응답을 목표로 하여 우선 텍스트에 의한 한국어 질의 응답 시스템을 호텔 예약을 텍스트로 하여 시험적으로 학계와 공동으로 구성하였으며 계속해서 의미 해석에 의한 대화이해의 고도화에 초점을 맞추어 성능 향상을 꾀하고 있다.

음성 및 언어 처리 연구의 기반이 되는 음성DB 및 텍스트 DB의 중요성이 날로 커져 가고 있으며 본격적인 연구를 위해서는 초기에 꼭 확보되어야 한다.^[11] 음성DB의 경우, ETRI에서 일찍부터 중요성을 인식하고 준비하여 왔으나 소모되는 인력 예산 등의 문제로 본격적인 구축이 어려웠다. 최근에 지금까지의 제한된 연구의 중간 결과물로서 단음절, 단독 및 연결 숫자음 등 일부 데이터들을 필요로 하는 학계에 시험적으로 배포한 바 있다.(93년 5월 현재 14개 대학 연구팀) 한편, 국가적인 공통 음성 데이터의 요구가 높아감에 따라 이에 대해 본격적으로 검토 되어 일부 결과가 CDROM화를 목표로 편집중에 있어 93년 중으로는 배포가 가능할 것으로 보인다. 한편, 텍스트DB의 경우는 우선 다양한 장르를 대상으로 한 300만 어절 수준의 자료를 입력하였고, 이를 tagging 하기 위한 도구들을 개발하고 있다.^[12]

V. 요소기술별 현황 및 과제

여기서는 자동통역전화의 실현에 필수적인 각 요소 기술들의 현황을 개략적으로 정리하고 해결해야할 과제들을 생각해 봄으로써, 아직 국내에서는 초기단계인 이분야의 연구 활성화에 다소나마 참고가 되었으면 한다.^[14]

자동통역전화시스템의 실현을 위해서는 그림6과 같은 요소기술들이 필요하다.

이들 기술의 현황 및 과제를 좀 더 자세히 살펴보

면 다음과 같다.

1. 대어휘 연속음성인식

실용가능한 자동 번역 전화 시스템 개발을 위해서는 연속 발성 방식에 의한 음성인식의 연구가 필요하다. 대어휘 연속 음성인식 기술을 고도화하기 위해서는 음소 특징의 안정적인 추출, 조음결합 문제의 해결 등 여러 가지 음향처리 레벨의 연구과제가 있으며, 문법, 의미, 문맥등의 언어적 지식을 지식원으로 유효하게 이용하는 기본기술을 축적하여야 한다.

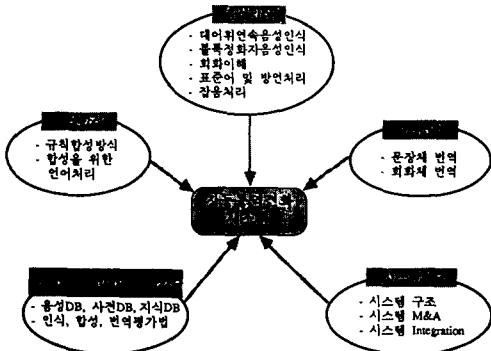


그림 6. 자동통역전화시스템의 구성을 위한 요소기술

1) 음향레벨 처리기술

연속음성 중 음소 등의 음성단위를 검출, 인식하는 기술로서, 현재 1000단어 수준의 인식은 단어단위로 가능하나 그 이상은 단어단위 인식으로는 한계가 있다. 따라서 음소단위의 인식방법이 쓰이고 있으며 음소단위 인식시에는 인접 음소간의 조음결합의 문제를 해결해야 한다.

○ 과제

- 유사음소를 구별할 수 있는 파라미터 추출
- 다량의 음성 데이터에 의한 조음결합의 모델화
- 특징 및 규칙탐색을 위한 전문가적 수법
- 피치 등 운율정보의 이용

2) 언어레벨 처리기술

음향 처리 레벨에서의 음소 인식의 정도에는 한계가 있기 때문에, 그 결과를 단지 조합하는 것만으로는, 단어와 문장의 인식율이 불충분하다. 이 때문에 단어, 문법, 의미 등의 언어적 지식과, 문맥등의 정보를 인식에 이용하는 기술을 개발해야 한다.

○ 과제

- 한국어 문법 확립 및 이용기술
- 지식처리에 의한 일반적 의미 이해기술을 이용한 인식
- 문맥이해기술 확립 및 인식에 적용

2. 불특정화자 음성인식

자동통역 전화 시스템의 실현, 보급을 위해서는 인식 대상범위가 불특정 다수화자를 대상으로 하여야 한다. 따라서 불특정화자 음성 인식에 관한 연구가 필수적이다. 불특정 화자 음성인식 기술을 확립하기 위해서는 음성의 개인성 특징의 정규화 및 화자 적응 등의 학습 방식이 유효한 수단이 된다.

1) 음성의 개인성 및 정규화법 연구

음성의 개인성은 성도 길이의 차이에 의한 정상 스펙트럼 차이와 조음결합의 정도의 차이 등으로 나타난다고 알려져 있다. 이러한 음성신호에 포함되는 개인성의 특징을 추출하고, 이를 정규화하기 위한 연구가 필요하다.

○ 과제

- 자 모음패턴의 개인별 분포 검토 및 개인성 요인 연구
- 화자변동에 강한 파라미터, 개인성추출에 적합한 파라미터 연구

2) 화자적응기술

정규화만으로는 모든 화자에 충분히 대응할 수 없다. 그래서 미리 발성한 비교적 소량의 음성, 또는 시스템 사용중의 입력음성을 기초로 하여 시스템을 화자에 적응 시키는 연구를 한다.

○ 과제

- 화자적응을 위해 적합한 단어나 문장 연구
- 음소에 대한 효율적 멀티템플레이트 작성법
- 유사특징을 갖는 화자 그룹간 식별기술
- 학습이 필요 없는 화자적응기술

3. 회화 이해기술

전화로 이야기하는 회화음성을 정확하게 인식하기 위해서는 단순한 연속음성 인식 및 이해의 틀을 넘어 회화 특유의 상황, 현상 메카니즘을 해명하고, 이에 대응하는 인식방법을 구축할 필요가 있다.

○ 과제

- 회화문 코퍼스 수집 및 모델화
- 일상 회화를 기술 가능한 문법 연구
- 회화모델에 의한 화제추정

- 운율정보 추출 및 이용기술
- 비정형문 (정정문, 불완전문, 오류문 등) 처리법

4. 잡음처리 및 회선특성

자동통역 전화가 사용되는 장소에는 기계음, 주위 사람의 말소리 등의 소음이 존재하고, 또, 통신회선에서는 혼신, 유도잡음 등이 섞여서, S/N비(신호대잡음비)가 나쁘다. 또, 전화기에서는 송화기의 주파수 특성의 변동, 고조파 등이 있고, 통신회선에서는 대역제한, 전송주파수특성의 변동등이 있어 음성파형을 변형시키거나 음성인식 성능을 떨어뜨리는 요인이 된다.

자동통역 전화의 실용화를 위해서는 이들 요인의 제거 또는 그 영향을 줄이는 기술개발이 필요하다.

1) 잡음처리

S/N이 열악한 신호속에서 원하는 신호만을 추출하는 방법과 S/N이 열악한 상태에서도 인식성능을 높게 유지하는 인식수법을 개발한다. 또, 입력장치에서 주위 잡음을 받지 않게 하는 방법, 송화기의 특징 변화나 그 영향을 줄이는 방법을 개발한다.

○ 과제

- 잡음하에서 음성 경계 검출법
- 잡음 학습에 의한 잡음 억압기술
- 복수 마이크에 의한 잡음 줄이는 연구
- 주변음성, 전화호출음 등 비정상적 잡음 제어법 연구

2) 회선특성

내선은 큰 문제가 없으나, 총체적인 회선상황은 어려운 상황이다.

○ 과제

- 전화망 전송특성 종합연구(상대국 포함)
- 회선특성 자동측정 및 등화방법 연구
- 회선혼신 대책 (주된 음성추적, 화자식별에 의한 음성분리 등)

5. 표준어의 지역차 및 방언연구

대상으로 하는 사용자가 전국적으로 다양한 언어를 사용하므로 표준어의 지역에 따른 차이(같은 어휘를 사용하더라도 억양의 미묘한 차이 등), 방언 등의 연구가 필요하다. 국어학계에서의 방언 연구가 있으나 제한적이며 표준어의 지역차 연구를 실험음성학적으로 시도한 경우는 거의 없다.

○ 과제

- 전국 각 지역 화자의 표준어 및 방언 데이터 수집 분석
- 표준어 음성의 표준적 특징 및 특징의 허용범위 연구
- 방언을 인식시스템에 적응시키는 연구

6. 문자언어 번역기술

한일간 기술문헌등의 번역시스템이 개발된바 있고 일한번역시스템은 일부 상품화된 것이 있으나 그 성능은 제한적이다.

○ 과제

- 문법연구 (문체, 문형, 문맥 등 포함)
- 지식베이스에 의한 의미처리 연구
- 문맥해석연구
- 텍스트 생성연구

7. 음성언어 번역기술

회화문의 경우는 외국의 경우도 아직 기초단계이며 국내에서의 본격적인 연구는 아직 없다.

○ 과제

- 회화문 번역기술연구
- 화자와 인터랙티브한 번역방식연구
- 회화문 생성기술 연구

8. 규칙합성방식

자동통역 전화 시스템의 음성합성부에는 임의의 문장음성을 적절한 운율과 자연스러운 음질로 합성하는 기술이 필요하다.

1) 음성합성방식

음성의 규칙합성 기술 가운데 합성의 기본단위를 연결해 연속 음성파형을 얻는 기본방식을 연구하는 것으로, 국내외 각국에서 자국어 합성 시스템이 상품화 되어 있으나 기계적이며 자연성이 부족한 제한된 음질이어서 이에 대한 개선이 진행중이다. 또, 복수의 임의 언어를 합성하는 시스템이 필요하다.

○ 과제

- 고품질 파형 합성방식 연구
- 복수언어 합성이 가능한 합성단위 연구
- 자연성을 포함한 합성음의 객관적 평가법

2) 음성합성규칙

임의의 연속음성의 합성에 필요한 합성규칙에 대한 연구로써, 합성규칙에는 합성 단위의 연결, 변형에 관한 음운규칙(분절 규칙)과 억양 등에 관한 운율규

칙(초분절 규칙)이 있다. 현재 수준으로는 음소변형의 체계적 규칙화가 어렵고, 언어정보로부터 자동으로 억양정보를 추출하는 능력이 아직 부족하며, 부분적인 운율연구로 음질이 아직 제한적이다. 특히 우리말의 경우, 대량의 실험음성학적 자료가 절대적으로 부족하다.

○ 과제

- 연속음성의 음운변형 및 조음결합규칙
- 음소 전이특성해명 및 합성단위 연결법
- 회화음성의 운율규칙
- 언어정보로부터 자동 운율규칙 추출 연구
- 의도, 스타일등 비언어적 정보에 의한 운율 실현
- 언어에 의존하지 않는 다언어 합성법

3) 합성을 위한 언어처리

문자와 문법, 의미정보로 음성을 합성하는 과정에서 필요한 언어처리 방법에 대한 연구로써, 불완전한 언어처리로 운율정보 자동추출이 아직 미흡하고 의미, 문맥해석 등 고도 언어처리기술은 아직 적용하지 못하고 있다.

○ 과제

- 언어정보로부터 운율, 음운레벨의 정보 도출
- 합성을 위한 언어처리 시스템 구성

4) 합성음의 음색제어

합성음성의 음색을 제어할 수 있는 음성합성 기술의 연구로, 성별음색, 연령별음색 제어가 미흡하고 합성을 위한 개인성 연구는 국내에서는 아직 없다. 또, 기본적인 합성음의 품질이 불충분하여 음색제어 연구에 장애가 되고 있다.

○ 과제

- 성별, 연령별 음색 합성법
- 개인성 실현규칙 연구
- 음색 제어를 위한 일반수법 연구

9. 음성 및 사전 데이터 베이스

음성인식 기술 및 음성합성 기술의 확립에는 음성 파형과 음성에 대한 언어내용의 대응관계를 해명하기 위해 그 기초가 되는 음성 데이터가 필요하다. 자동 번역 전화 시스템의 실현에는 음성 데이터를 데이터 베이스로서 조직화한 음성 데이터 베이스가 꼭 필요하다. 또, 번역에 필요한 상식(세계지식) 및 전문지식을 개념과 그 상호간 관계로 받아들여 기술한 지식 데이터 베이스의 구축방법을 확립해야 한다. 아울러 회화문의 번역에 필요한 언어사전의 구성, 구축 방법

을 확립해야 한다.

1) 음성 데이터베이스

연구 및 평가에 필요한 한국어 및 상대국어의 음성 데이터 베이스를 작성한다. 음성 데이터 베이스는 통역전화의 개개의 연구에 있어 기본적 도구인 동시에 개발내용 평가에 기준을 부여하는 것으로 연구 초기 단계부터 필요하며, 또, 그 축적은 상당히 긴 시간을 요하는 일이기 때문에 조기에 집중적으로 연구개발할 필요가 있다. 음성 데이터 베이스는 각 단계의 연구에 선행하여 준비할 필요가 있으며, 조기에 집중적으로 작성할 필요가 있다. 또, 연구용 데이터 베이스는 연구 목적에 맞는 것을 작성할 필요가 있으며, 다음 단계 연구를 진행함에 있어서의 목표로서 서브테마마다 목표를 설정한다. 또, 연구 목적에 입각해 각각 음성 데이터 베이스의 형식을 결정할 필요가 있다.

○ 과제

- 대어휘 연속음성인식용
- 잡음처리 및 회선특성연구용
- 규칙합성용

2) 음성DB 작성, 관리, 지원 시스템

간단한 파일 구조를 주로 이용하고 있으며, 관계형 DB를 이용한 경우도 있으나 음성DB의 속성을 감안한 DB구조가 필요하다.

○ 과제

- 음성DB 검색시스템
- 데이터 추가변경 등 버전관리 시스템
- 음성 수록 편집 시스템의 표준사양
- 데이터 저장 포맷 연구

3) 언어사전DB

기계번역을 위한 사전이 번역시스템마다 소규모로 작성되어 있으나 제한적이다.

○ 과제

- 사전항목, 표기법 설정
- 사전DB 구축 및 음성DB와 통합

4) 지식베이스

지식의 표현에 의한 지식정보의 축적이 필요하나 DB화할 대상 데이터의 선정이 아직 어렵고 개념의 표현 및 체계화가 필요하다.

○ 과제

- 개념의 선택기준 및 기술방법 설정
- 개념체계의 구축
- 지식베이스 구축
- 지식 및 언어베이스 관리 및 지원시스템 구축

10. 시스템 구조연구

1) 기본시스템의 구조

음성인식, 합성, 기계번역의 각 요소기술을 유기적으로 통합화하여 자동통역 전화 시스템의 기능을 실현시키기 위한 시스템의 구성 및 이때 필요한 대용량 고속 하드웨어에 요구되는 성능에 관해 검토할 필요가 있다.

○ 연구과제

- 기본 시스템을 서브시스템으로 최적 분할하는 방법과 상호 인터페이스 방식을 확립
- 각 서브시스템의 기능, 성능을 실현하는데 필요한 처리내용, 속도와 병렬처리 가능성등을 검토 및, 하드웨어의 대략적인 형식의 결정

(2) 토털 시스템의 구조

번역 통신을 위한 토털 시스템으로써 구체적으로는 네트워크 구성과 기능분담, 통신절차, 과금방식 등에 관해 검토한다. 또, 시스템 운용에 필요한 각종사전, 데이터 베이스의 구성에 관해서도 검토한다.

○ 연구과제

- 국제 전화망 내에서 번역 통신 시스템의 최적 배치와 서브 시스템의 기능 분담
- 각종 통신 서비스의 형태를 고려한 후에 번역 통신의 절차, 과금방식 등에 관하여 검토
- 번역 통신 시스템 운용에 필요한 각종 사전, 데이터 베이스에 대하여 분산화, 계층화 등의 검토 및 관리, 운용 방식에 관한 검토

VI. 결론

지금까지 자동통역전화에 관한 국내외의 연구동향 및 그 실현을 위한 과제에 대해서 정리해 보았다. 국제간의 상담, 여행, 통신 등이 빈번해진 요즘, 그 어느 때 보다도 통역서비스의 필요성이 커지고 있다. 최근에는 국제통화에 통역자를 통한 통역지원서비스가 실시되고 있고, 여행자들을 위한 다국어 번역지원기가 판매되고 있는가 하면, PC통신을 통한 번역지원 서비스도 이루어지기 시작하고 있다. 그러나 아직까지는 만족 할만한 수준에 이르렀다고 보기는 어렵다. 이를 위해서는 본격적이고 체계적인 음성, 언어 정보처리 관련의 공학적 연구 뿐만 아니라 언어학, 음성학, 심리학, 철학, 논리학 등 다양한 학문간의

학제적 연구가 필요할 것이다. 또한 궁극적으로 언어 정보에 의한 통신이라는 관점에서 "컴퓨터커뮤니케이션 사이언스" 라는 관점의 연구도 주장되고 있다.^[14] 아울러 통역은 반드시 상대국이 있게 마련이므로 각기 자국어에 대한 전문성을 가지고 국제협력에 의해 풀어야 할 숙제를 많이 안고 있어서 연구의 국제화가 필요하다. 이러한 자동통역 서비스를 통신망을 통해 제공하는 것은 지능화(Intelligent), 복합화(Multimedia), 개인화(Personal), 인간화(Human) 등 "IMPH"^[15] 를 지향하는 최근의 정보통신기술 개발 방향에서도 불 때도 인간위주의 서비스를 제공하는 핵심기술의 하나로 그 파급효과가 클 것으로 기대된다. 이 분야의 지속적인 지원과 관련 연구자들의 공동 노력이 필요한 시점이다.

參 考 文 獻

[1] Akira Kurematsu, "Speech and Language Processing for Automatic Telephone Interpretation" *J. of IEICE, Japan* vol.75 no.10 (92.10.)

[2] Stentiford, Steer, "Machine translation of speech" *British Telecom, Technology Journal* vol. 6 no. 2 (April 1988)

[3] Preprints of International Symposium on Prospects and Problems of Interpreting Telephony The Telecommunications Advancement Foundation of Japan (April 12, 1986)

[4] 이용주, "일본 ATR연구소의 첨단연구 소개" 한국전자통신연구소 전자통신동향분석 (93. 1.)

[5] 우정성 자동번역전화시스템 개발추진협의회 보고서 일본우정성 자동번역전화시스템 개발추진협의회 (86.7.)

[6] 김정호, "ATR의 자동통역전화 연구현황" 한국정보과학회 1993년도 춘계학술발표회 특강 요약집 (93.4.)

[7] Waibel, Jain, et al. "JANUS: A speech to speech translation using connectionist and symbolic processing strategies"

Proc. ICASSP-91(April 1991)

[8] Sagayama, et al. "An Experimental Interpreting Telephone System 'ASURA'" ASJ Spring Meeting 3-4-17. (1993-3)

[9] 자동통역전화를 위한 요소기술 개발(I)(II) 한국전자통신연구소 연구보고서 (1991, 1992)

[10] 이종락 "한국통신의 자동통역연구현황" 제9회 음성통신 및 신호처리 워크샵 논문집, 한국음향학회(92.8)

[11] 정유현, 이용주, "음성 데이터 베이스의 연구 동향 및 전망" 한국음향학회지 제10권 제4호 (1991)

[12] 이용주, 임연자 외, "ETRI의 음성 및 텍스트 데이터 베이스의 구축 현황" 제1회 ETRI 음

성, 언어 및 음향정보처리 워크샵 논문집 (1993.4.)

[13] 김경태, 이용주, 정유현, "자동통역전화시스템 구성을 위한 검토" 대한전자공학회 통신, 교환 연구회 발표집(87.11)

[14] S. Chiba, "Interpreting Telephony" J. of IEICE , Japan vol.70 no.4 (April 1987)

[15] 양승택, "정보통신 기술개발 추진방향" 한국정보과학회 1993년도 춘계학술발표회 특강 요약 집 (93.4.)

筆者紹介



李 勇 柱

1954年 1月 17日生

1976年 2月 고려대학교 전자공학과(학사)

1987年 8月 고려대학교 대학원 전자공학과(석사)

1992年 8月 고려대학교 대학원 전자공학과(박사)

1976年 3月 ~ 1980年 7月 공군통신장교 근무

1985年 7月 ~ 1986年 2月 일본 토호쿠대학 응용정보학 연구센터(연구생)

1980年 8月 ~ 현재 한국전자통신연구소 자동통역연구실 실장(책임연구원)

주관심분야: 음성, 언어 및 음향정보처리, HCI