

음성패턴인식 인터랙티브 콘텐츠 개발

Interactive content development of voice pattern recognition

나종원*

Jong-Won Na*

요 약

언어 학습 콘텐츠에서 공통적으로 가질 수 있는 문제점들을 분석하고 문제점에 대하여 음성 패턴인식기술을 적용하여 기존의 문제점을 해결하였다. 언어 학습 콘텐츠의 첫 번째 문제점은 온라인 학습 자세이다. 수업 진행은 되었지만 다른 웹 페이지를 열어 게임을 하는 등 학생들의 집중력은 떨어졌다. 두 번째 문제점은 Speaking 학습 과정을 만들었지만 실제로 따라 읽는지 판단할 수가 없었다. 세 번째 문제점은 학습 관리 시스템에 의한 기계적 진행이 아니라 선생님들의 평가에 의해 잘하는 학생들과 못하는 학생간의 학습 진행에 차이를 둘 필요가 생겼다. 마지막으로 가장 큰 문제는 기존에 만들어 놓은 콘텐츠들은 그대로 유지되면서 위의 문제들을 해결할 수 있어야 했다. 이러한 배경 하에 음성 패턴인식기술은 말하기 학습 전용 학습 프로그램으로 학습 진행을 위한 음성인식은 물론 학습 자체를 위한 음성인식 기능들을 모두 가지고 있으며 인식 절차에 사용된 학습자의 발화 데이터를 원하는 형태의 오디오 파일로 변경하여 서버의 특정 위치로 전송하거나 SQL서버에 등록할 수도 있으며, 또한 컴포넌트이기 때문에 그 어떠한 시스템이나 프로그램이라도 모두 적용 가능하고 이미 만들어진 콘텐츠 전체를 손상시키지 않고 쉽게 삽입하여 새로운 기능들을 사용할 수 있었다. 본 논문으로 교육 방식을 보다 인터랙티브하게 바꾸어 적극적인 수업참여가 되도록 기여하였다.

Abstract

Voice pattern recognition technology to solve the problems of the existing problems and common issues that you may have in language learning content analysis. This is the first problem of language-learning content, online learning posture. Game open another web page through the lesson, but the concentration of the students fell. Have not been able to determine the second issue according Speaking has made the learning process actually reads. Third got a problem with the mechanical process by a learning management system, as well by the teacher in the evaluation of students and students who are learning progress between the difference in the two. Finally, the biggest problem, while maintaining their existing content made to be able to solve the above problem. Speaking learning dedicated learning programs under this background, voice pattern recognition technology learning process for speech recognition and voice recognition capabilities for learning itself has been used in the recognition process the data of the learner's utterance as an audio file of the desired change to a transfer to a specific location of the server or SQL server may be easily inserted into any system or program, any and all applicable content that has already been created without damaging the entire component because the new features were available. Contributed to this paper, active participation in class more interactive teaching methods to change.

Key words : Voice pattern recognition, Interactive, contents, flash, engine, jsfl, swf

I. 서 론

본 논문은 IT 기반 하에서 다양한 네트워크 및 디

* (주)유비바다 미래기술연구소장
· 제1저자 (First Author) : 나종원
· 투고일자 : 2012년 9월 25일
· 심사(수정)일자 : 2012년 9월 25일 (수정일자 : 2012년 10월 23일)
· 게재일자 : 2012년 10월 30일

바이스를 이용하여 학습자에게 언제 어디서든 환경에 관계없이 제공되어질 수 있는 초·중급 영어 교육생 대상의 음성패턴인식 인터랙티브 영어교육 콘텐츠 설계 및 구현을 통해 학습자의 요구를 향상시켰다. 또한, 개발하고자하는 속성의 논문이 전무하여 연구의 어려움이 있었다.

교육과정을 학습자가 조절하여 사용자의 학습 역량에 맞춘 개인화된 학습을 학습자 스스로 진행할 수 있는 학습자 위주의 교육 콘텐츠를 개발하였다.

인터넷 관련 기술의 발달은 점차 교실 수업의 비중을 온라인 수업으로 옮기도록 하고 있으며 의사소통 능력·말하기 능력을 중요시 하는 언어 교육 분야에서도 음성·언어 기술의 발달과 함께 온라인 말하기 학습 시대를 열고 있다[3].

이에 음성패턴 인식기술은 온라인 언어 학습에 필요한 전문 기능들을 하나로 묶어, 일반 웹 페이지는 물론 플래시 같은 다른 응용 프로그램에서도 활용 가능 하도록 제작된 기술이다.

음성패턴 인식기술은 학습 목적에 맞는 응용 프로그램인 플래시에서 다음과 같은 기능을 적용하기 위한 목적이 있다[4]. 그 기능으로는 마이크를 통해 들어오는 학습자의 음성 데이터를 문장·단어·음소·의미 등의 학습 목적에 적합한 모드로 인식하여 학습에 적용하기 위해 음성으로 변환하여 학습자에게 들려주는 기능과 그 이외의 언어 학습에 유용한 기능들을 말한다.

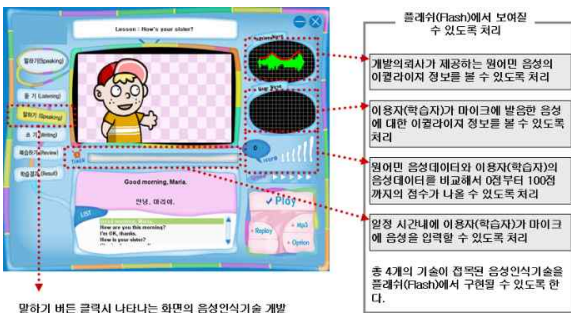


그림 1. 음성패턴인식 인터랙티브 콘텐츠 예시화면
Fig 1. Voice pattern recognition illustrative screen interactive content

II. 관련연구

2-1 디지털콘텐츠 활용을 위한 정부정책

교육과학기술부는 2010년 1월 창의적 산지식을 제공하고, 학습자 친화적인 미래형 교과서 보급을 주요 골자고 한 “2010년 교과서 선진화 방안“을 확정 발표하고 디지털 콘텐츠 활용 의지를 천명한바 있다. 또한 “스마트러닝 세계 4대 강국으로 도약“이라는 목표를 설정하고 지경부, 교과부, 행안부, 문화부, 고용부 및 그 외 민간위촉위원을 위원으로 이터닝산업발전위원회를 발족 제2차 이터닝산업 발전 및 활성화 기본계획을 확정하여 추진 중이다[1].



그림 2. 2011. 2. 28 정부발표 이터닝산업 계획[1]
Fig 2. Government announced the E-Learning Industry Plan[1]

2-2 시장 활성화에 부흥

웹 기반에서 디지털콘텐츠 IT 기술 발전을 통해 디지털콘텐츠 기반 이터닝(e-Learning)산업이 매년 성장해 관련 산업규모는 2조원을 넘어서고 사업체수도 1500개가 넘는 것으로 나타났다[6].



그림 3. 정보통신진흥원 “2010년 이터닝산업 실태조사 결과[1]
Fig 3. “2010 E-Learning Industry Survey Results[1]

2011년 3월28일 지식경제부와 정보통신산업진흥원이 발표한 이터닝산업 실태조사 결과, 2010년 이터닝 시장규모는 2조2458억 원으로 전년대비 7.4% 성장했다. 부문별로는 서비스부문이 전년대비 8.9%성

장한 1조5129억 원으로 전체의 67.3%를 차지했으며 콘텐츠는 3.7%증가한 5090억 원, 솔루션은 6.0%증가한 2239억 원으로 각각 집계됐다[5].

국내 이러닝 산업계는 디지털콘텐츠 업체와 서비스업체가 함께 성장해 나가는 상생협력모델을 발굴하고 이러닝 콘텐츠, 솔루션, 서비스 등 이러닝 패키지의 해외 진출을 활성화할 수 있도록 수출 유망모델 발굴과 현지화 지원에 나설 계획이다[1][7].

III. 시스템의 구성

3-1 음성패턴인식 인터랙티브 콘텐츠 적용기술

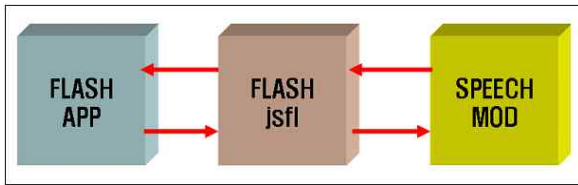


그림 4. 음성패턴인식 기술 구조
Fig 4. Voice pattern recognition technology structure

.jsfl 스크립트 파일과 엔진 dll 을 통해 통신하는 모듈이다. 플래시는 디렉트로 엔진을 접근 할 수는 없으며, jsfl 을 통하여 연결을 시도한다. 플래시의 형태는 swf 파일 포맷 형태를 유지한다. 필요에 따라 다양한 APP 로 이식이 가능하다.

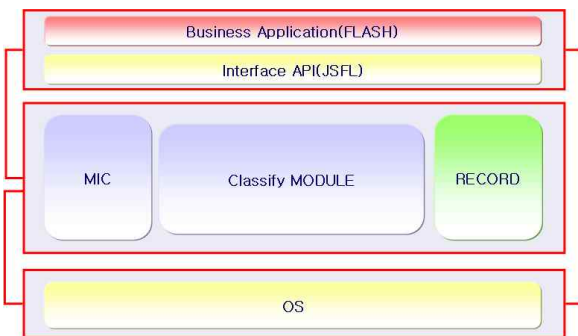


그림 5. 엔진 프레임워크
Fig 5. Engine Framework

상호작용을 위한 인터랙티브 콘텐츠 기술은 단일 콘텐츠와는 달리 난이도나 유형별로 한 가지 학습주제에 대해 여러 단계의 콘텐츠로 구성되어 있다.

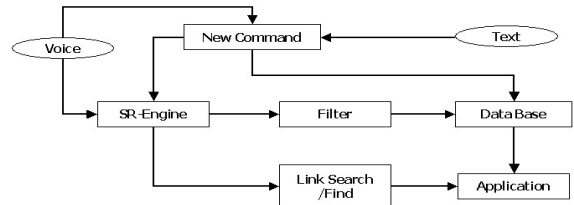
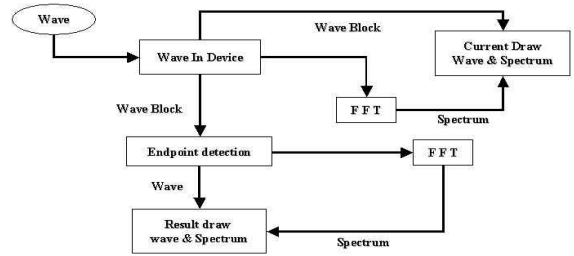


그림 6. 음성패턴인식 도식도
Fig 6. Voice pattern recognition schematic

학습자는 기초적인 문제 풀이 및 음성패턴인식 부터 시작하여 과제를 해결하였을 경우에는 보다 높은 난이도의 문제에 도전 할 수 있다. 과제해결에 실패하였을 경우에는 풀이과정을 학습하고 다시 문제를 해결하도록 반복학습 한다. 기존의 학습자 수준을 고려하지 않았던 학습 콘텐츠의 한계를 뛰어넘어 학습자의 학습능력수준에 대응되는 인터랙티브 콘텐츠를 학습자에게 제공하고자 하는 것이다.

캐릭터 액션 스크립트에서 사용자의 답안 및 음성 패턴인식 결과에 따라 피드백을 받아 조건을 분기하여 적절한 보충 설명과 함께 사용자의 수준에 맞는 콘텐츠를 선택하여 학습자에 제공하게 된다.

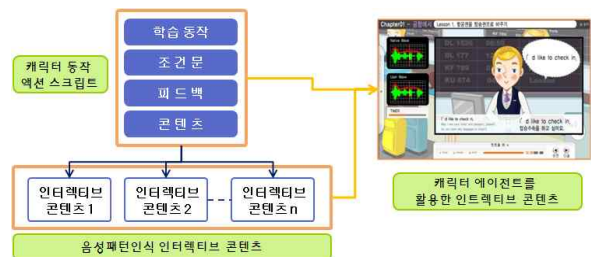


그림 7. 인터랙티브 콘텐츠
Fig 7. Interactive content

강의 방법 중 캐릭터 에이전트의 다양한 애니메이션의 활용은 학습자의 관심을 사로잡기에 효과적인 요소이다. 그러므로 복잡하고 자칫 지루해지기 쉬운 영어교육 과정에서 캐릭터의 사용은 학습자의 학습

흥미를 유도 할 수 있을 것이다. 그러나 캐릭터의 애니메이션은 미리 정해진 몇 가지 행위로 한정되어 있기 때문에 일정시간동안 반복되는 동작을 접한 후 학습의 흥미가 떨어질 수 있다.

이를 보완하기 위해 다양한 종류의 캐릭터 에이전트 모델을 추가하여 학습자는 다양한 상황에 따라 새로운 캐릭터 모델 및 애니메이션을 접하면서 학습 할 수 있다.

3-2 설계단계

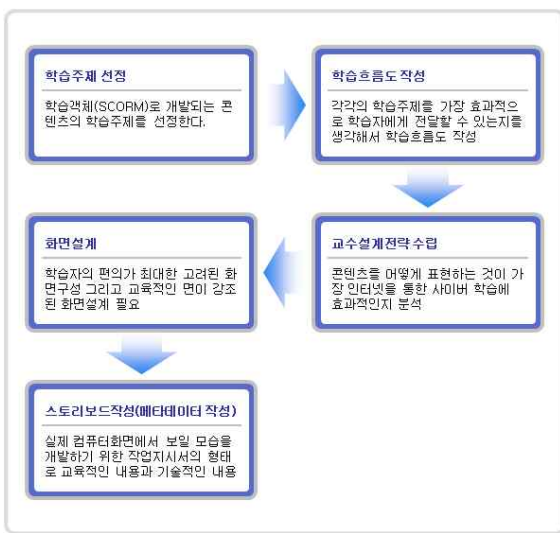


그림 8. 설계단계 흐름도
Fig 8. Design phase flowchart

설계단계는 선정된 과정에 대한 원고를 인터넷을 통한 멀티미디어 콘텐츠로 제작하기 위한 기본구성 단계이다. 이 단계에서는 학습객체로 개발되는 콘텐츠의 학습주제를 선정하게 되고 각 학습주제를 가장 효과적으로 전달할 수 있는 학습흐름도를 결정한다.

학습흐름이 결정되면 어떻게 내용을 구분하여 전달하고 어떻게 표현하는 것이 가장 인터넷을 통한 사이버 학습에 효과적인지를 분석하여 이를 위한 전략을 세운다. 이러한 교수설계전략은 각 학습 흐름 별로 작성하여야 하며, 스토리보드의 형태로 표현되게 한다. 또한 이러한 스토리보드를 작성할 때 공유유통을 위한 메타데이터를 결정하고 작성한다.

차시 : 현재 페이지의 차시를 입력한다.
 학습제목 : 해당 차시의 제목을 입력한다.
 작성자 : 스토리보드를 작성한 담당자 이름을 입력한다.
 학습메뉴 : 차시에서 해당하는 SCO 명을 입력한다.
 디렉토리명 : 차시의 디렉토리명을 입력한다.
 개발유형 : 화면구성 내용에 대한 개발유형을 입력한다.

화면구성
학습 콘텐츠의 화면상에 보여지는 애니메이션이나 액션 버튼 등의 구성을 스케치 한다.

나레이션
이 화면에서 필요한 나레이션을 학습자가 이해하기 쉽도록 학습을 설명을 한다.

부가 정보항
콘텐츠에 사용되는 스크린, 이미지, 나레이션, 플래시애니메이션, 스타일시트 그리고 제작의 도움이 필요한 참고문서나 파일의 위치를 입력한다.

F/C
SCO별 각 Depth별로 구분해서 현재 설명하고 있는 부분이 어디인지를 쉽게 알 수 있도록 한다.

그림 9. 설계단계 산출물
Fig 9. Design phase deliverables

3-3 제작단계

제작단계는 스토리보드에 따라 웹디자인, 멀티미디어제작 등 미디어요소를 제작하여 최종 콘텐츠를 제작하는 단계이다. 반드시 하나 이상의 프로토타입을 개발하고 스토리보드의 내용과 실제 콘텐츠와의 적합성을 파악하여야 하며, 기본적인 디자인의 질을 확보하여야 한다. 이러한 프로토타입의 개발이 완료되고 확정되면 각 멀티미디어콘텐츠가 개발된다.

구현 기술인 음성패턴인식 인터랙티브 기술을 포함한 기술개발을 한다.

그리고 이러한 콘텐츠는 반드시 전체적인 교수설계에 충실하여 제작한다. 이렇게 콘텐츠 제작이 이루어지면 최종 개발된 콘텐츠에 대한 검수를 한다.

마지막으로 콘텐츠 제작/검수가 완료되면 SCORM (Sharable Content Object Reference Model) 표준화 패키징 작업을 하게 된다.

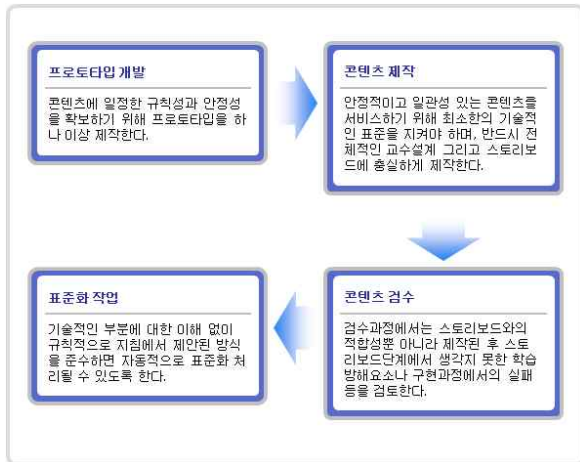


그림 10. 제작단계 흐름도
Fig 10. Production step flowchart

여 입모양을 보면 사용자(학습자)가 쉽게 발음을 교정 할 수 있도록 한다.

IV. 테스트 및 완료

음성패턴인식 인터랙티브 콘텐츠를 제작해서 테스트 단계로 웹기반 전자교육에 대한 표준규격 SCORM (Sharable Content Object Reference Model) 적용에 문제가 없는지, 학습모형이 제대로 되었는지, Site 탑재 시 오류 및 버그 발생은 없는지 등을 테스트 하는 단계이다.

Lesson 1. 향공권을 탐승권으로 바꾸기



Lesson 1. 향공권을 탐승권으로 바꾸기



그림 11. 음성패턴인식 인터랙티브 적용 학습평가 인터페이스

Fig 11. Learning Assessment applies interactive voice pattern recognition interface



그림 12. 테스트단계 흐름도
Fig 12. Test step flowchart

콘텐츠의 표준화란 특정 목적을 위해 일정한 규약을 공통으로 적용하는 것이다. 개발된 콘텐츠가 서로 검색이 가능하고 필요에 따라 새로 개발하여 교체하거나 같은 시도 또는 다른 시도의 콘텐츠를 혼합하여 재구성하고자 하는 경우 반드시 표준화되어 있어야 한다[2]. 또한 SCORM의 표준화는 학습자의 학습활동을 추적함으로써 학습효과를 극대화하고 학습자 개인에 맞는 맞춤형 수준별 학습이 가능하도록 하는 장점이 있다.

학습평가에서는 온라인 교육콘텐츠에서 주기적인 음성(패턴)인식기술이 접목된다.

원어민의 음성과 사용자(학습자)의 음성을 비교 분석하여 0부터 100점 까지 점수를 배점하면 배점된 점수가 기준(60점 이상)이 되지 않으면 피드백을 하고 다시 문제를 풀 수 있도록 한다.

비드백은 원어민의 입모양을 동영상으로 처리하

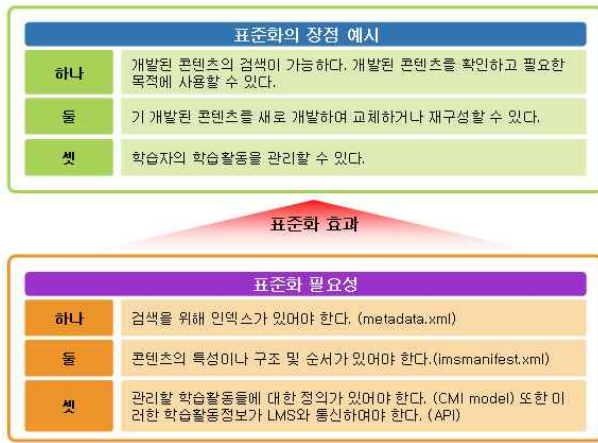


그림 13. 표준화의 필요성[2]
Fig 13. Necessity of standardization[2]

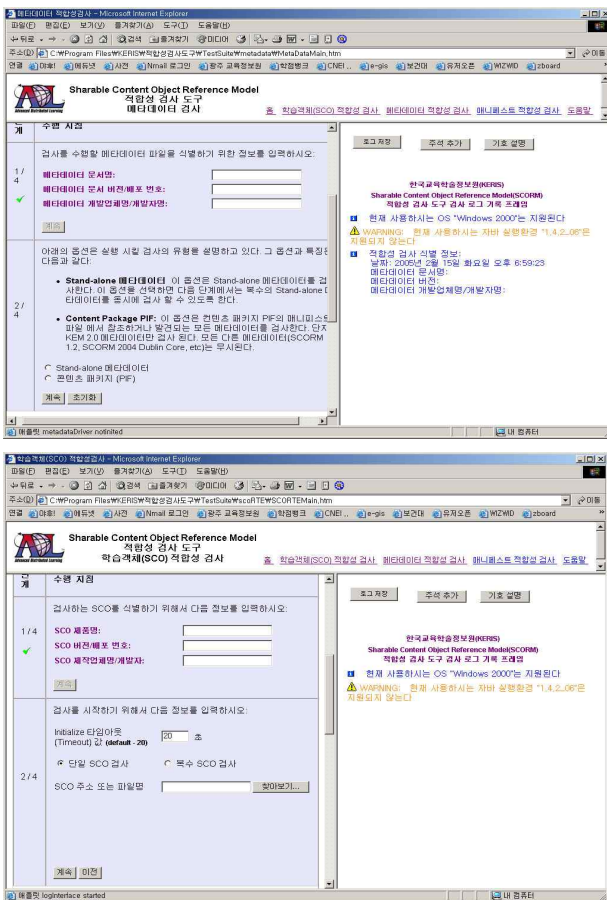


그림 14. 테스트 및 표준화 적합성 도구
Fig 14. Conformance testing and standardization tools

Lesson 1. 향 공권을 탑승권으로 바꾸기



Lesson 1. 향 공권을 탑승권으로 바꾸기



그림 15. 음성패턴인식 인터랙티브 적용 화면
Fig 15. Interactive applies screen voice pattern recognition

IV. 결 론

언어 학습 콘텐츠의 문제점은 수업 진행 중 다른 웹 페이지를 열어 게임을 하는 등 수업에 대한 집중력이 떨어지며, Speaking 학습 과정을 만들었지만 실제로 따라 읽는지 판단할 수가 없었고, 학습 관리 시스템에 의한 기계적 진행이 아니라 선생님들의 평가에 의해 잘하는 학생들과 못하는 학생간의 학습 진행에 차이를 둘 필요가 생겼다.

음성 패턴인식기술은 학습 진행을 위한 음성인식은 물론 학습 자체를 위한 음성인식 기능들을 모두 가지고 있으며, 인식 절차에 사용된 학습자의 발화 데이터를 원하는 형태의 오디오 파일로 변경하여 서버의 특정 위치로 전송하거나 SQL서버에 등록할 수도 있다. 또한 컴포넌트이기 때문에 그 어떠한 시스템이나 프로그램이라도 모두 적용 가능하고 이미 만들어진 콘텐츠 전체를 손상시키지 않고 쉽게 삽입하여 새로운 기능들을 사용할 수 있었다.

본 연구는 언어 학습 콘텐츠에서 공통적으로 가질 수 있는 문제점들을 분석하고 문제점에 대하여 음성 패턴인식기술을 적용하여 기존의 인식으로 해결하지

못했던 문제점을 해결하였고 교육과정을 학습자가 조절하여 사용자의 학습 역량에 맞춘 개인화된 학습을 학습자 스스로 진행할 수 있는 학습자 위주의 교육 콘텐츠를 개발하였다. 또한, 인터랙티브하게 바뀌어 적극적인 수업참여가 되도록 유도하였다.

참 고 문 헌

- [1] <http://www.mke.go.kr/news/bodo/bodoView.jsp?pCtx=1&seq=66542>
- [2] 장재경, “3D 애니메이션 콘텐츠의 SCORM 기반 표준화 전략”, *한국콘텐츠학회*, 2006. 2
- [3] 송정영, “음성인식을 위한 새로운 포맷트랙킹 알고리즘의 제안과 평가”, *인터넷정보학회논문지*, 제13권, 제4호, 2002. 6
- [4] 안중영, “자동차 ECU제어를 위한 음성인식 패턴매칭레벨에 관한 연구”, *한국인터넷방송통신학회 논문지*, 제10권, 제1호, 2010. 8
- [5] 김종훈, “음성인식을 이용한 상황정보 기반의 스마트 홈 개인화 서비스”, *한국콘텐츠학회 논문지*, 제9권, 제1호, 2009
- [6] 백동광, “화자 종속 알고리즘을 이용한 음성 인식 시스템 구현 화자 종속 알고리즘을 이용한 음성 인식 시스템 구현” *대한전자공학회 학술대회 논문집* 제4권, 제6호, 2004. 10.
- [7] 안중영, “모델적응 HMM을 이용한 모바일환경에서의 음성인식에 관한 연구” *한국인터넷방송통신학회 논문지*, 제11권, 제3호, 2011. 10.

나 종 원 (羅鍾元)



2003년 전남대학교 소프트웨어공학
협동과정 공학석사

2009년 전남대학교 소프트웨어공학
협동과정 공학박사

2005년~2007년 정보통신부 IT국내
교수요원, 동강대학 정보통신과
초빙전임강사

2009년~2010년 광주대학교 산학협력단 전담교수

2010년~현재 (주)유비바다 미래기술연구소장

관심분야 : 얼굴인식, 음성인식, 웹솔루션