

음성녹음 기반의 실감형 어학시스템 콘텐츠 개발

Development of tangible language content system based on voice recording

나종원*

Jong-Won Na*

요 약

기존의 어학시스템 콘텐츠의 문제점은 수업에 대한 집중력 저하와 실제 학습여부를 판단할 수 없었고, 수업 진행자의 평가로 많은 결정이 되었다. 이로 인해 음성녹음 기반에 유비쿼터스 기술과 가상현실 기술을 조합하여 강의실에 프로젝터를 설치하고 각 강의실의 RFID 리더기 부착과 학생의 RFID 태그 부착된 학생증을 통해 영어 등급에 해당하는 콘텐츠를 학습한다. 3차원 영상 콘텐츠로 가상 현실상의 외국인과의 질의응답을 음성녹음 기술을 이용하여 발음과 억양을 동시에 체크 하여 레벨 패스 또는 레벨 실패를 판가름 한다. 이렇게 구성된 시스템은 학생 교육 데이터를 중앙 서버에 DB로 저장후 피드백 과정을 통해 정보를 제공하게 된다. 본 연구로 어학 콘텐츠에서 공통적으로 가질 수 있는 문제점들을 분석하고 문제점에 대하여 음성 녹음기술을 적용하여 기존의 어학콘텐츠에서 해결하지 못했던 문제점을 해결하였고 레벨위주의 수업이 가능하였다. 또한, 흥미위주의 적극적인 수업참여가 되도록 기여하였다.

Abstract

Learning a lesson about poor concentration and problems of the existing content, the system of language which could not be determined, Many teachers' assessment decision was made. As a result, voice recording based on the combination of ubiquitous technology and virtual reality technology, and install the projector in a classroom Through the learning content corresponding grade English student ID card attached RFID reader in each classroom, and students of RFID tags attached. In reality of the virtual three-dimensional image content foreigners and question-and-answer using the voice recording technology at the same time check the pronunciation and intonation level passes or level failure judged. Student education data to a central server system is configured to do so after saving to the DB through a feedback process, which provides information. Analysis of the issues that can have a common language content in the present study and Problem for voice recording technology to solve the problem and did not solve the existing language in the content level based classes.

Key words : Voice recording, RFID, E-Learning, Contents

I. 서 론

본 논문은 음성녹음 기반 하에서 네트워크 및 디바이스를 이용하여 학습자에게 실감형 어학시스템

* 미래기술연구소 소장(Mi-rae Technical Institute)

· 제1저자 (First Author) : 나종원(Jong-Won Na, tel : +82-11-614-5612 , email : najwon@empal.com)

· 접수일자 : 2013년 4월 5일 · 심사(수정)일자 : 2013년 4월 5일 (수정일자 : 2013년 4월 24일) · 게재일자 : 2013년 4월 30일

<http://dx.doi.org/10.12673/jkoni.2013.17.2.234>

콘텐츠를 제공하여 학습자의 요구를 향상시키기 위함이다.

영어캠프를 위해 해외로 나가던 학생들을 흡수하기 위해 시작된 국내 영어 마을 사업과 각 시·도의 영어캠프가 성과를 거두고 있다. 하지만 그 수요층이 한정 되어있고, 고비용과 사교육을 조장하는 것이 현실이다. 이에 유비쿼터스 기술과 가상현실 기술을 조합하여 몇 개의 소형 강의실에 일반 프로젝트터 또는 3차원 영상을 제공해 줄 수 있는 프로젝터를 설치하고 각 강의실 문에는 RFID 리더기를 부착하여 학생이 RFID 태그가 부착 되어 있는 학생증을 가지고 강의실에 입장하였을 경우 시스템에서는 카드를 인식하여 그 학생의 영어 등급에 해당하는 콘텐츠를 보여주게 된다.



그림 1. RFID 흐름도
Fig 1. RFID flowchart

이때 해당 콘텐츠는 3차원 영상으로 공항, 은행, 편의점 등 외국과 동일한 환경의 배경과 더불어 가상 현실상의 외국인이 학생에게 영어로 질문 하며 학생은 영어로 질문에 대한 답을 한다. 답변에 대한 내용은 음성녹음 기술을 이용하여 발음과 억양을 동시에 체크 하여 레벨 패스 또는 실패를 판가름 한다. 이렇게 구성된 시스템은 학생 교육 데이터를 중앙 서버에 DB로 저장후 스마트폰을 통해 수강한 수업자료 또는 다음레벨 수강 자료를 전송하여 향후 다시 공부 하고자 하였을 경우에 피드백 역할을 하는 등의 정보를 제공하게 된다.

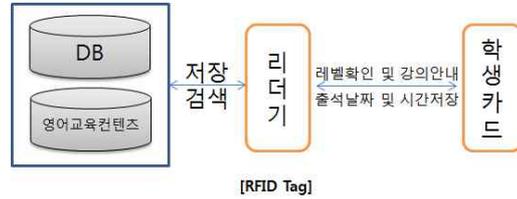


그림 2. RFID-태그
Fig 2. RFID-Tag

II. 관련연구

2-1 다양한 기반의 어학 콘텐츠

■ 유비쿼터스(Ubiquitous)기반의 어학 콘텐츠

“언제, 어디서나, 누구든지” 다양한 네트워크를 이용하여 교육 서비스를 이용하고, 보다 풍요로운 학습 환경을 구현하는 것으로 학습자가 외출 또는 이동 중일 때도 콘텐츠의 접속이 가능하며, ADSL, FTTH(Fiber to the Home)등의 브로드밴드회선, 제3세대 휴대전화, 무선랜, Bluetooth 등 다양한 네트워크에서도 연결이 가능하며, 지금까지 PC를 사용할 수 없어서 교육 콘텐츠를 이용할 수 없는 다양한 계층을 포함하여 폭넓은 사람들이 교육을 받을 수 있는 어학 콘텐츠를 말한다.

■ 광대역통합망(BcN, Broadband convergence Network)에 연동하여 제공할 수 있는 다채널 교육 콘텐츠

정부의 IT839 전략의 중점 추진 과제로 진행 중인 사업인 광대역통합망은 유비쿼터스 환경을 구축위한 유무선 통합, 통신과 방송 융합, 음성 및 데이터의 통합을 통한 차세대 광대역 통신망(FTTH, beyond 3G 등)을 기반으로 개인화 및 주문화 된 고품질 양방향 멀티미디어 콘텐츠 서비스 제공을 위한 통로가 개설되고 있다. 이러한 네트워크 통합을 대비한 차별화된 광대역 멀티미디어 콘텐츠 제작 및 서비스를 제공할 수 있는 다양한 포맷의 어학 콘텐츠이다.

■ Rich Media 어학 콘텐츠

사용자와의 상호 작용이 가능한 멀티미디어로 Java, Shockwave, VRML, SMIL 언어를 이용하여 교육

콘텐츠에서 학습자에게 다양한 상호 인터랙티브한 정보를 교환하면서 학습자들이 개개인의 수준에 맞춘 교육 시간표와 학습 능력에 맞는 교육 프로그램을 학습자 스스로 만들어 갈 수 있는 상호 인터랙티브한 멀티미디어 어학 콘텐츠이다.

■ 음성패턴인식기술을 적용한 어학 콘텐츠

온라인 어학에 필요한 전문 기능들을 하나로 묶어 일반 웹 페이지는 물론 플래시 같은 다른 응용 프로그램에서도 활용 가능 하도록 제작된 기술이다.

음성패턴 인식기술은 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 [3].

2-2 시장 활성화에 부흥

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



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

2011년 3월28일 지식경제부와 정보통신산업진흥원이 발표한 이러닝산업 실태조사 결과, 2010년 이러닝 시장규모는 2조2458억 원으로 전년대비 7.4% 성장했다. 부문별로는 서비스부문이 전년대비 8.9% 성장한 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 시스템의 적용기술

시스템의 적용은 RFI/D, 디스플레이 및 다양한 소프트웨어로 구성하였다.

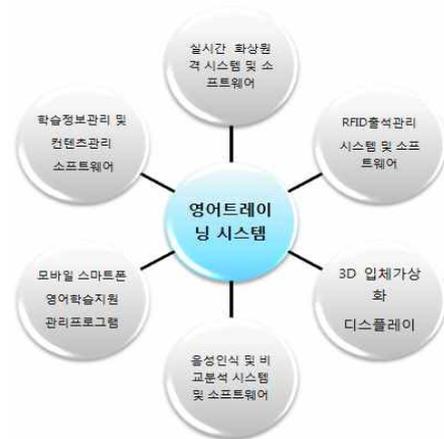


그림 4. 시스템의 적용기술
Fig 4. Apply the technology of the system

3-2 과제해결 및 과제해결수단

각 사용자들이 강의실을 이용해서 자신의 영어 학습 등급에 맞는 다양한 주제의 영어 회화 상황을 직접 경험해 보도록 하여, 접근성과 현실감이 뛰어난 환경 속에서 체계적으로 영어 회화 실력을 키워나갈 수 있도록 하는 가상공간에서의 영어 학습 서비스 시스템을 제공하는데 그 목적이 있다.

해결수단으로는 영어 학습 서비스 시스템은 각 강의실에 일반프로젝터 또는 3차원 영상을 제공해 줄 수 있는 프로젝트가 설치되어 시각적/청각적 인터페이스를 제공하는 클라이언트 PC와 서버 PC의 통신망을 통해 상기 각 클라이언트와 연동하여 각 사용자들의 영어 학습 등급에 따라 가상공간에서 수행할 일

정 임무를 안내하고 상기 각 클라이언트와의 사이에서 상기 안내된 미션을 진행하며 각 사용자들의 미션 수행 결과에 따라 해당 사용자의 영어 학습 등급을 관리하는 서버를 포함하여 이루어진다.

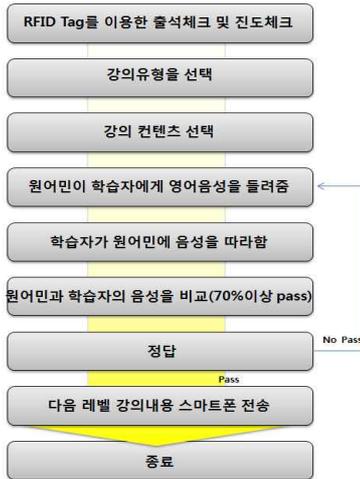


그림 5. 과제해결을 위한 프로세스
Fig 5. Challenges for process

3-3 시스템 구성도

음성녹음 기반의 실감형 어학시스템 콘텐츠 개발을 위한 시스템 구성도이다. 본 시스템은 3차원 가상 현실을 구현하기 위해 3D가상 캐릭터 또는 3D카메라를 이용하여 촬영하고 그 콘텐츠를 3D프로젝터를 이용하여 화면에 디스플레이하여 학습자가 실감형으로 느낄수 있다.

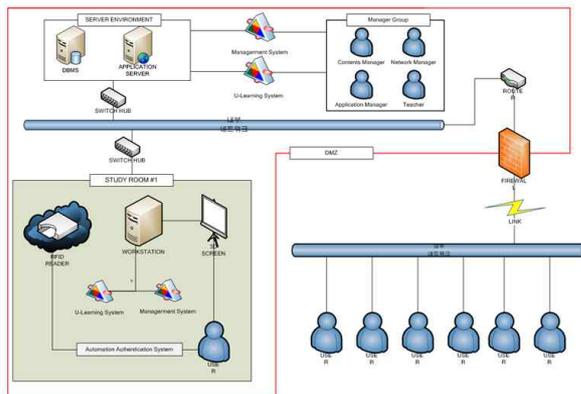


그림 6. 시스템 구성도
Fig 6. System Configuration

3-4 시스템 이용 조감도

3차원 영상을 제공의 프로젝터를 설치하고 각 강의실 문에는 RFID 리더기를 부착하여 학생이 RFID 태그가 부착 되어 있는 학생증을 가지고 강의실에 입장하였을 경우 시스템에서는 카드를 인식 하여 그 학생의 영어 등급에 해당하는 콘텐츠를 보여 주게 된다.



그림 7. 시스템 이용 조감도
Fig 7. System using a bird's eye view

3-5 음성비교 프로세스

음성녹음을 통해 음성을 비교하여 발음과 억양을 동시에 체크할 수 있는 음성비교 프로세스이다.

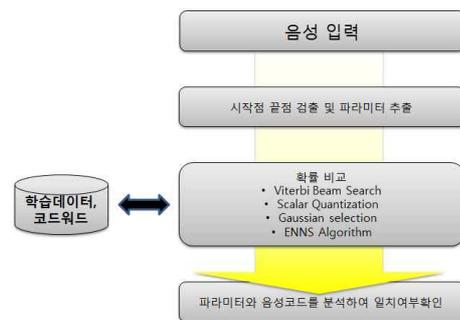


그림 8. 음성비교 프로세스
Fig 8. Voice comparison process

IV. 테스트 및 완료

음성녹음 기반의 실감형 어학시스템 콘텐츠를 제작해서 테스트 단계로 웹기반 전자교육에 대한 표준 규격 SCORM (Sharable Content Object Reference Model) 적용에 문제가 없는지, 학습모형이 제대로 되

있는지, Site 탑재 시 오류 및 버그 발생은 없는지 등을 테스트 하는 단계이다. 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].

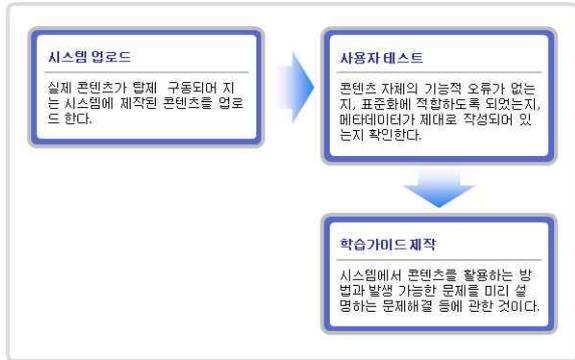


그림 9. 테스트단계 흐름도[3]
Fig 9. Test step flowchart

또한 SCORM의 표준화는 학습자의 학습활동을 추적함으로써 학습효과를 극대화하고 학습자 개인에 맞는 맞춤형 수준별 학습이 가능하도록 하는 장점이 있다.

표준화의 장점 예시	
하나	개발된 콘텐츠의 검색이 가능하다. 개발된 콘텐츠를 확인하고 필요한 목적에 사용할 수 있다.
둘	기 개발된 콘텐츠를 새로 개발하여 교체하거나 재구성할 수 있다.
셋	학습자의 학습활동을 관리할 수 있다.

표준화 효과	
표준화 필요성	
하나	검색을 위해 인덱스가 있어야 한다. (metadata.xml)
둘	콘텐츠의 특성이나 구조 및 순서가 있어야 한다. (imsmanifest.xml)
셋	관리할 학습활동들에 대한 정의가 있어야 한다. (CMI model) 또한 이러한 학습활동정보가 LMS와 통신하여야 한다. (API)

그림 10. 표준화의 필요성[2]
Fig 10. Necessity of standardization[2]

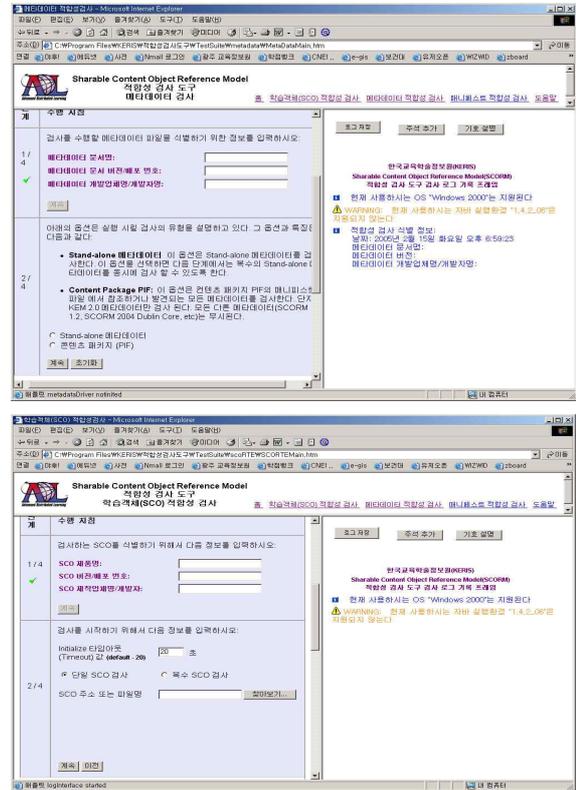


그림 11. 테스트 및 표준화 적합성 도구[3]
Fig 11. Conformance testing and standardization tools

강의콘텐츠는 3차원 영상으로 은행, 공항, 식당 등 외국과 동일한 환경의 배경과 더불어 가상 현실상의 외국인이 학생에게 영어로 질문 하며 학생은 영어로 질문에 대한 답을 한다. 답변에 대한 내용은 음성녹음 기술을 이용하여 발음과 억양을 동시에 체크 하여 레벨 패스 또는 레벨 실패를 판가름 한다.

강의콘텐츠

1. 은행

Teller : May I help you sir?
Hojin : I'd like to cash my traveler's check.
Teller : Okay, how much do you want to cash?
Hojin : About 500 dollars.
Teller : I need your ID and traveler's check.
Hojin : All right. Here you are.
Teller : Thank you, sir. Just one moment, please.
Hojin : Okay.

2. 공항입국장

A: Please hand me the customs declaration form.
B: Here it is.
A: Open your baggage, please.
B: I have only personal effects.
A: You can go out through green line.
B: Thank you.

A: I'd like to see that pen, please.
B: You mean this one?
A: No, the other one in the brown case.
B: Oh, this one. Here.
A: May I try it?
B: Sure.

3. 쇼핑

4. 식당

Guest: That was very nice, thank you.
Waiter: Would you like a sweet, madam?
Guest: Yes, but nothing too filling. What do you suggest?
Waiter: An ice, perhaps? The lemon sorbet, for instance?
Guest: No, not that. I'll have the Creme Caramel.
A: May I try it?
Waiter: Creme Caramel. And would you like coffee afterwards?
Guest: Yes, please.

그림 12. 어학시스템 강의 콘텐츠
Fig 12. Language systems lecture content

IV. 결 론

기존의 어학시스템 콘텐츠의 문제점은 수업에 대한 집중력 저하와 실제로 학습하는지를 판단할 수가 없었고, 수업 진행자의 평가로 많은 결정이 되었다.

RFID 리더기 부착과 학생의 RFID 태그 부착된 학생증을 통해 영어 등급에 해당하는 콘텐츠를 학습한다. 가상 현실상의 외국인과의 질의응답을 음성 인식 기술을 이용하여 발음과 억양을 동시에 체크 하여 레벨 패스 또는 실패를 판가름 한다. 이렇게 구성된 시스템은 학생 교육 데이터를 중앙 서버에 DB로 저장 후 피드백 과정을 통해 정보를 제공하게 된다.

본 연구는 어학 콘텐츠에서 공통적으로 가질 수 있는 문제점들을 분석하고 문제점에 대하여 음성 녹음 기술을 적용하여 기존의 어학콘텐츠에서 해결하지 못했던 문제점을 해결하였고 레벨위주의 수업이 가능하였다. 또한, 흥미위주의 적극적인 수업참여가 되도록 기여하였다.

Reference

[1] <http://www.mke.go.kr/news/bodo/bodoView.jsp?pCtx=1&seq=66542>
 [2] J.k.jang , "Standardization Strategy on 3D Animation Contents" *Journal of Korea Contents Association*, Vol.14 No 2, 2006
 [3] J.W.Na "Interactive content development of voice pattern recognition", *The journal of Korea Navigation Institute*, Vol.16 No 5, 2012
 [4] J.Y.An, " A Study on Voice Recognition Pattern matching level for Vehicle ECU control ", *Journal of Advanced Smart Convergence*, Vol.10 No 1, 2010
 [5] J.H.Kim, "Smart Home Personalization Service based on Context Information using Speech Recognition ", *Journal of Korea Contents Association* , Vol.9 No 11, 2009
 [6] d.G.Baik, " Speaker dependent speech recognition system using the algorithm implementation " *Journal of the Institute of Electronics Engineers of Korea Conference*

Proceedings, Vol.2004 No 6, 2004

[7] J.Y.An, "A study on Voice Recognition using Model Adaptation HMM for Mobile Environment " *Journal of Advanced Smart Convergence*, Vol 11.No 3, 2011. 10.

나 종 원 (Jong-Won Na)



2003년 전남대학교 소프트웨어공학 협동과정 이학석사
 2009년 전남대학교 소프트웨어공학 협동과정 공학박사
 2005년~2007년 정보통신부 IT국내 교수요원, 동강대학교 정보통신과 전임강사

2009년~2010년 광주대학교 산학협력단 전담교수
 2010년~현재 미래기술연구소 소장

관심분야 : 얼굴인식, 음성인식, 웹솔루션, 정보시스템