

전자프로그램 가이드를 위한 음성 인터페이스

김한수⁰ 황인준
아주대학교 정보통신전문대학원 정보통신공학과
{kfreeman, ehwang}@ajou.ac.kr

VoiceEPG: Speech Interface for Electronic Program Guide

Hansu Kim⁰ Eunjung Hwang
Graduate School of Information and Communication, Ajou University

요약

최근 디지털 TV 방송의 활성화에 힘입어 수많은 채널을 통한 TV 프로그램 방송이 가능하게 되었다. 이로 인해 디지털 TV 시청자들은 신문 또는 TV 가이드와 같은 기존 인쇄매체를 통해 자신이 원하는 TV 프로그램 스케줄을 얻기가 사실상 매우 어렵게 되었다. 이와 같은 문제점을 해결하기 위해 디지털 TV 환경에서는 전자 프로그램 가이드(EPG: Electronic Program Guide)를 제공한다. 현재 제공되고 있는 EPG 서비스들은 대개 디지털 TV 화면 또는 각 방송사 웹 사이트 그리고 이동 단말기 등을 통해서 서비스 되고 있다. 대부분의 기존 연구들은 EPG 정보를 화면상에 시각적으로 제공하는 측면에만 초점을 두고 있다. 하지만 실질적으로 사용자 입장에서는 원하는 방송 프로그램의 스케줄 정보를 찾기 위해서 수백 채널에 달하는 방송 프로그램에 대한 정보를 일일이 검색하는 것은 매우 힘든 일이다. 게다가 사용자가 원하는 키워드를 직접 입력하는 방식 또한 사용자들 매우 번거롭게 한다. 따라서 본 논문에서는 EPG 서비스 방식에 VoiceXML 관련 기술을 접목하여 이동 단말기상에서 간단한 음성입력을 통해 EPG 서비스를 제공받을 수 있는 음성 인터페이스를 제안한다.

1. 서론

디지털 TV 방송이 시청자들에게 멀티미디어 정보 제공의 필수적인 역할을 하게 됨에 따라 수많은 채널을 통한 방송 프로그램의 신설, 추가, 변경, 삭제가 빈번해 지고 있다. 따라서, 시청자들은 신문 또는 TV 가이드를 통해 방송 프로그램에 대한 스케줄 정보 획득이 매우 어렵게 되었다.

이같은 문제점을 보완 하고자 디지털 TV 환경에서는 EPG를 제공한다. EPG는 앞으로 방송될 프로그램의 제목, 방송 시간, 프로그램 검색(채널별, 시간대별, 날짜별, 키워드), 그리고 프로그램의 상세요약 등의 정보들을 제공한다. 이러한 EPG 서비스는 계속적인 방송 채널의 신설이 예상됨에 따라 사용자들에게 매우 중요한 수단이 된다.

또한, 최근 이동 단말기 사용의 급속한 수요 증가와 맞물려 언제, 어디서나 정보를 얻으려는 사용자들의 요구가 크게 증가하고 있다. 하지만, 이동 단말기 상에서 시각적으로 EPG 서비스를 검색하기에는 화면의 제약성과 텍스트 입력을 통한 검색 방식의 불편함을 가진다.

따라서 본 논문에서는 이동 단말기 상에서 VoiceXML 관련 기술을 적용하여 기존 검색 방식의 단점을 극복하고 시각적으로 장에 놓인 상황에서도 EPG 서비스를 검색할 수 있는 음성 인터페이스를 설계 및 구현 하였다.

대표적인 VoiceXML 관련 기술과 EPG의 특징과 이슈들을 간략히 살펴 보면 다음과 같다. VoiceXML의 주된 목표는 유선 전화기 또는 이동 단말기 상에서 간단한 버튼 조작

이나 음성 입력을 통하여 VoiceXML Interpreter에 접속하고 검색 내용 또한 음성으로 제공 받는 것이다. 이같은 서비스를 제공하기 위해서 VoiceXML은 음성인식(Speech Recognition), 음성변환(Text To Speech), 버튼입력 감지(DTMF Input), 음성입력의 저장(Recording of Spoken Input)과 같은 기능들을 포함하고 있다[1].

또한, 최근 디지털 TV와 대화형 TV 기술의 발전으로 인하여 EPG에 대한 관심과 이를 이용하려는 연구들이 크게 증가하고 있다. EPG는 TV 프로그램 방송시간과 내용, 출연자에 대한 정보 등을 데이터베이스로 구축하고 이를 기반으로 개인의 취향을 반영하는 맞춤형 프로그램 가이드를 생성하고 제공하는 서비스이다. 따라서 EPG를 이용하여 사용자에게 적합한 맞춤형 서비스를 생성하고 제공하기 위한 다각적 측면의 연구들이 활발히 진행되고 있다[2,3,4].

본 논문의 구성은 다음과 같다. 2 장에서는 제안된 VoiceEPG 시스템의 전체적인 구조와 각 구성요소들의 상호 역할을 기술한다. 3 장에서는 EPG XML 스키마 구조를 알아보고, 4 장에서는 컨버터를 통한 VoiceXML 문서의 개발 과정을 기술한다. 마지막으로 5 장에서는 결론에 대해 논의 한다.

2. VoiceEPG 시스템

본 논문에서 제안하는 시스템은 XML 형태로 제공되는 EPG 정보를 EPG XML 데이터베이스에 저장한다. 내부적으로 EPG XML 문서내 엘리먼트 콘텐츠를 추출하여 VoiceXML

문서들을 생성하기 위해 오프라인상에서 VoiceEPG Converter를 통한 지속적인 변환 과정이 이루어진다. 그림 1은 전체적인 시스템 구조를 보여준다.

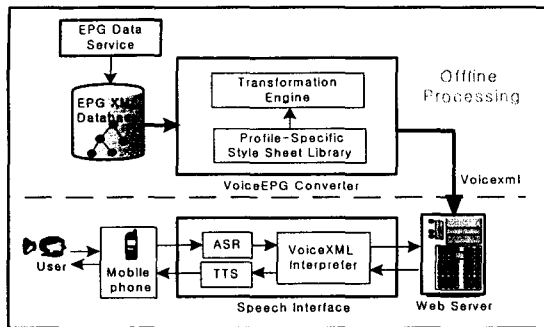


그림 1. VoiceEPG 시스템 구조

VoiceEPG 시스템은 네 개의 구성 요소들로 이루어져 있다.

- **EPG XML DataBase:** 각 방송사별로 제공되는 EPG 서비스를 저장하고 정보가 갱신될 때마다 VoiceEPG Converter로 전송하는 역할을 수행한다.
- **Speech Interface:** 내부적으로 음성인식 엔진은 사용자의 음성을 인식하고, 사용자 요청 질의 처리 제어권을 VoiceXML Interpreter로 넘기는 역할을 한다. VoiceXML Interpreter는 각 사용자의 요청을 모니터링 하면서 음성인식 엔진으로부터 전송된 이벤트에 해당하는 VoiceXML 문서를 웹 서버로부터 제공받아 파싱하는 역할을 한다. 또한 VoiceXML Interpreter는 음성변환 엔진을 제어하여 파싱한 VoiceXML 문서를 통해 음성 대화를 진행해 나간다.
- **VoiceEPG Converter:** 오프라인 상에서 EPG XML 데이터베이스로부터 새로운 EPG 서비스 정보가 추가될 때마다 갱신 정보를 받아서 VoiceXML 문서로 변환하는 역할을 수행하게 된다. 내부적으로 프로파일 스타일시트 라이브러리에는 EPG 서비스를 제공하기 위한 다양한 XSLT 시나리오 파일들이 저장되어 있다. 또한 변환엔진은 EPG XML 문서들과 XSLT 시나리오 파일들을 결합하여 결과 VoiceXML 문서들을 생성한다.
- **Web Server:** 웹 서버는 VoiceEPG Converter를 통해 변환된 VoiceXML 문서들을 저장하는 역할을 하고 HTTP 프로토콜을 통해 VoiceXML 인터프리터가 요청한 EPG 서비스 정보를 포함하는 VoiceXML 문서를 전송하는 역할을 가진다.

3. EPG XML 스키마 분석

본 논문에서는 EPG XML 엘리먼트내의 콘텐츠를 추출하여 VoiceXML 문서를 생성하기 위해 Eureka-147 디지털 오디오 방송 스키마 스펙을 분석하였다[5]. 이 스키마 스펙은 세계의 독립적인 스키마들을 포함한다. 즉, i) 일반 데이터 타입에 관한 스키마 부분, ii) 스케줄에 관한 스키마 부분, iii) 서비스 정보에 관한 스키마들로 이루어져 있다. 그림 2는 이 중에서 실험을 위해 사용된 스케줄에 관한 스키마의 계층 구조를 나타낸다. epg는 루트 엘리먼트로서 schedule, programmeGroups, programme 그리고 alternateSource의 하위 엘리먼트들을 포함한다.

programme 엘리먼트는 특정 방송 프로그램의 제목, 서비스되는 방송국, 프로그램에 대한 짧은 요약, 상세 요약 서비스, 장르 정보, 키워드 정보, 연관된 프로그램의 목록 정보, 이벤트 정보를 포함하게 된다.

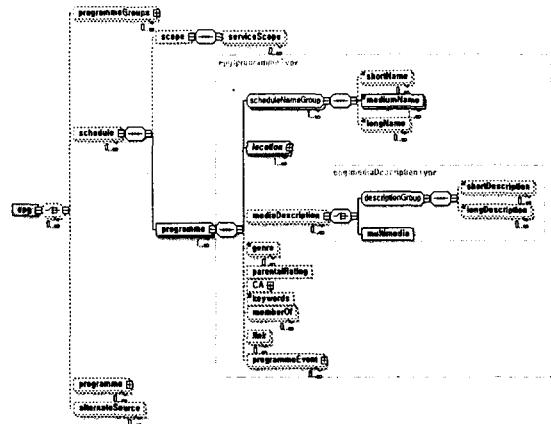


그림 2. EPG XML 스케줄 스키마 구조

4. 실험

본 실험에서는 EPG XML를 VoiceXML로 효과적인 변환하기 위해 EPG XML 문서내 엘리먼트의 콘텐츠를 추출하여 VoiceXML 문서를 생성하는 컨버터를 구현하였다. 그림 3은 VoiceEPG 시스템에서 가장 중요한 역할을 하는 VoiceEPG Converter의 구현 프로토타입을 보여준다. 왼쪽 영역은 원본 EPG XML 문서 트리 구조를 나타내고 오른쪽 영역은 변환된 결과 VoiceXML 문서를 나타낸다. VoiceEPG Converter는 사용자 인터페이스를 위해 Java Swing과 JAXP(Java API for XML)[6]를 사용하고 내부적인 처리를 위해 XSLT 변환과정을 거친다.

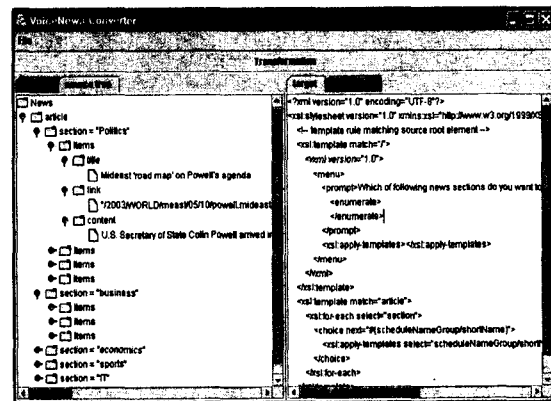


그림 3. VoiceEPG Converter

구현된 VoiceEPG Converter의 내부적인 XSLT 변환 처리 과정을 영화 요약 EPG XML 서비스의 예를 통해 설명하면 다음과 같다. 그림 4는 EPG XML 스펙에 기반한 원본 EPG XML 문서를 나타낸다.

```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="EPG2VoiceXML.xml"?>
<epg xmlns:epg="http://www.worlddab.org/schemas/epg"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.worlddab.org/schemas/epg
  epgSchedule_10.xsd" system="DAB">
<schedule>
<programme id="crid://www.kbs.co.kr; dab/BC81123456"
  version="1" recommendation="yes" bitrate="160">
  <scheduleNameGroup>
  <shortNamexml:lang="en"> Title of a movie ...</shortName>
  <mediaDescription>
  <shortDescription xml:lang="en"> Contents of a movie ...
  </shortDescription>
  </mediaDescription>
  .....
  .....
</programme>
</schedule>
</epg>
```

그림 4. 원본 EPG XML 문서

그림 5는 VoiceXML 문서를 생성하기 위한 XSLT 템플릿 문서를 나타낸다. 또한 그림 6는 원본 EPG XML 문서와 XSLT 템플릿 문서를 VoiceEPG Converter의 변환엔진을 통해 생성한 VoiceXML 문서를 나타낸다.

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/
XSL/Transform" xmlns:fo="http://www.w3.org/1999/XSL/Format">
  <xsl:template match="/">
  <vxml version="1.0">
  <menu> <prompt>Which of following do you choose?
  <enumerate> </prompt>
  <xsl:apply-templates/>
  </menu>
  </vxml>
  </xsl:template>
  <xsl:template match="schedule">
  <xsl:for-each select="programme"> <choice next="{#schedule
  NameGroup/shortName}">
  <xsl:apply-templates select="scheduleNameGroup/shortName">
  </choice>
  .....
  .....
  </xsl:for-each>
  </xsl:template>
</xsl:stylesheet>
```

그림 5. XSLT 템플릿 문서

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<vxml version="1.0">
<menu>
<prompt>
  Welcome to VoiceEpg movie services
</prompt>
</prompt>
<prompt>
  Please, select one of : <enumerate>
</prompt>
<choice next="#beautiful_mind">
  Beautiful Mind </choice>
<choice next="#man_in_black">Man in Black</choice>
<choice next="#titanic">Titanic</choice>
<choice next="#cinema_paradiso">
  Cinema Paradiso</choice>
<choice next="#rain_man">Rain Man</choice>
</menu>
</vxml>
```

그림 6. 결과 VoiceXML 문서

생성된 결과 VoiceXML 문서들을 테스트 하기 위해 Windows 2000 환경에서 VoiceXML 개발 툴 로서 널리 사용되고 있는 모토롤라 모바일 애플리케이션 개발자 키드[7]로 검증해 본 결과 효과적으로 서비스됨을 확인할 수 있었다. 결과적으로 그림 7과 같이 이동 단말기 사용자는 간단한 음성 입력으로 VoiceXML Interpreter와 음성 대화를 통해 EPG 서비스를 검색 할 수 있게 된다.

```
VoiceXML Interpreter: Welcome to VoiceEPG movies service!
Please Select one of: Beautiful Mind, Man
in Black, Titanic, Cinema Paradise, Psycho.

User: Titanic

VoiceXML Interpreter: This short is simply wonderful! Now, the
true story of the last moments of the Titanic
have been revealed... Do you continue
service?

User: Yes

VoiceXML Interpreter: Welcome to VoiceEPG movies service! ...
```

그림 7. 음성 대화의 예

5. 결론

현재 제공되는 EPG 서비스는 시각적인 측면에만 초점을 두고 있다. 본 논문에서는 이동 단말기 상에서의 화면 제약성과 텍스트 기반 검색 방식의 단점을 극복 하기 위해 음성으로 EPG 서비스를 제공하는 기법을 제안하였다. 이에 대한 세부적인 사항으로 EPG XML의 내부 구조를 파악하기 위해 Eureka-147 디지털 오디오 브로드캐스팅의 EPG XML 스키마 스펙을 분석하였다. 또한, EPG XML 문서의 엘리먼트 계층구조에서 콘텐츠를 추출하여 VoiceXML 문서를 생성하는 XSLT 기반 VoiceEPG Converter를 구현하였다. 그리고 이와 같은 결과 VoiceXML 문서들을 VoiceXML Interpreter 환경에서 테스트 하여 효과적으로 서비스됨을 확인할 수 있었다. 특히, 이러한 검색방식에 대한 연구는 일반인은 물론 시각적으로 정보를 획득하기 어려운 운전자나 시각장애인과 같은 사람들에게 큰 도움을 줄 것으로 예상된다.

참고 문헌

- [1] VoiceXML Forum, Voice eXtensible Markup Language VoiceXML, <http://www.w3c.org/Voice/>
- [2] J. Lang and M. Strick: Experiments in personal television: The future of EPG project, *Proceedings of IBC 2001, Amsterdam 2001(CD ROM edition)*
- [3] www.nds.com/solutions/electronic_program_guideprint.html
- [4] ETSI/EBU ICS 33.020:ETS 300 707-EPG , May 1997
- [5] European Telecommunication Standard ETS 300 401: "Radio Broadcasting Systems: Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers".
- [6] Sun Microsystems, Java API for XML-Processing <http://java.sun.com/xml/jaxp/index.html>
- [7] <http://developers.motorola.com/developers/>