

이동 방송 환경에 적합한 사용자 상호작용에 관한 연구

이봉호, 양규태, 임형수, 허남호

한국전자통신연구원

leebh@etri.re.kr

A Study on the User Interaction for Mobile Broadcast Context

Bongho Lee, KyuTae Yang, Hyeongsoo Lim, Namho Hur

*Electronics and Telecommunications Research Institute (ETRI)

요 약

이동 방송 환경에서 사용자가 특정 프로그램 또는 특정 서비스를 편리하게 수신할 수 있도록 하기 위해서는 고정 환경과 달리 적합한 사용자 상호작용이 고려되어야 한다. 본 논문에서는 디지털 라디오 환경에서 청취자에게 서비스를 편리하고 효과적으로 접근하고 수신할 수 있도록 하는 멀티모달 기반의 사용자 상호작용 기술에 대해서 소개하고자 한다. 이를 위해 디지털 라디오 환경 및 서비스 구조를 분석하고 이를 기반으로 사용자 상호작용이 필요한 사항을 도출하였다. 사용자 상호작용 엔진은 XML 마크업 언어를 기반으로 서비스, 프로그램 및 콘텐츠 수준의 상호작용이 가능하도록 구성하였다.

1. 서론

디지털 라디오인 DAB[1]는 아날로그 FM 과 같이 깨끗한 음질을 제공하는 라디오 방송으로 이동 수신에 최적화된 시스템이다. 디지털 라디오 방송은 라디오 프로그램인 오디오 서비스를 기본으로 제공하여 부가적으로 정지영상이나 텍스트를 이용하여 프로그램과 관련 정보를 제공하거나 멀티미디어 서비스 형태인 BWS(Broadcast Websites)나 BIFS(Binary Format for Scene) 기반의 데이터 서비스 제공이 가능한 시스템이다.

수신기 관점에서 보면, 키친 라디오와 같이 고정형, 이동형 및 차량형 수신기로 크게 구분할 수 있으며 이들 수신기의 사용자 상호작용은 기본적으로 그래픽에 기반하고 있다. 단문 문자 서비스에 해당하는 DLS(Dynamic Label Service)를 포함하여 프로그램 전체를 한눈에 보고 선택할 수 있는 EPG 에서부터 웹 페이지 형태로 다양한 정보를 제공할 수 있는 BWS 까지 그래픽 기반의 인터페이스로 화면 출력과 버튼을 포함하여 GUI 기반으로 navigation 하고 선택할 수 있는 기능을 지원하는 형태의 인터페이스로 구성되어 있다.

이러한 환경은 차량과 같은 이동 환경의 사용자에게는 다소 부족한 면이 있다. 이는 운전자의 주위를 방해할 뿐만 아니라 정보에 접근하고 이를 습득하는 과정이 편리하지 않다. 일 예로 운전 중에 현재 방송중인 프로그램 아이템의 정보에 접근하고자 할 경우 운전자는 DLS 프로그램을 구동해야 하며 원하는 정보가 화면에 출력될 때까지 주시해야 한다. 이를

개선하기 위한 DL plus 서비스의 경우에도, 세부 정보에 접근하기 위해서는 적어도 2 단계의 선택 과정이 필요하다. 이러한 상호작용은 차량 환경의 사용자에게 다소 부족한 면을 지니고 있다.

라디오 방송 즉 청취를 기반으로 한 라디오 서비스에서는 특히 그래픽 기반의 정보 접근은 다소 제약이 있어 이를 개선할 방안이 요구되고 있다. 이를 위해서 서비스, 프로그램 및 특정 정보로의 접근이 그래픽을 포함하여 다양한 방식으로 가능하다면 이동 방송 사용자에게는 수신 환경을 개선하는 효과가 있을 것으로 사료된다.

본 논문에서는 전술한 기능이 가능하도록 하기 위해 멀티모달 기반의 사용자 상호작용 기술을 연구하였다. 멀티모달 상호작용이란 단일 인스턴스(instance)에 하나 이상의 인터페이스 즉 모달(modal)을 이용하여 대상과 상호작용 하는 것을 의미한다. 이를 위해 본 논문에서는 XLM 마크업 언어 기반의 상호작용 엔진을 기반으로 디지털 라디오 방송에서 제공하는 서비스 및 프로그램 객체에 접근하여 나아가 각 객체내의 세부 내용을 멀티모달 형태로 안내 받는 방법을 기술하고자 한다.

2. 이동 방송 시스템 환경 및 서비스 구조

본 논문에서는 디지털라디오 방식 중 하나인 DAB(Digital Audio Broadcasting)를 목표 시스템으로 선정하고 사용자

상호작용 측면에서의 환경 및 구조를 분석하였다. DAB의 경우, 약 1.5Mbps 내에 다수의 오디오 프로그램과 관련 부가 서비스 데이터들이 다중화 되어 전송된다. 다중화된 서비스 및 프로그램에 대한 식별 및 접근을 위해 FIC(Fast Information Channel) 채널을 통해 MCI(Multiplex Configuration Information) 정보를 전송하고 MSC(Main Service Channel)을 통해 각 채널을 구성하는 프로그램 데이터를 다중화하여 전송하고 있다. 이에 대한 구조는 그림 1로 도시할 수 있으며 EPG에서는 이러한 구조를 고려하여 정보를 구성한다.[1]

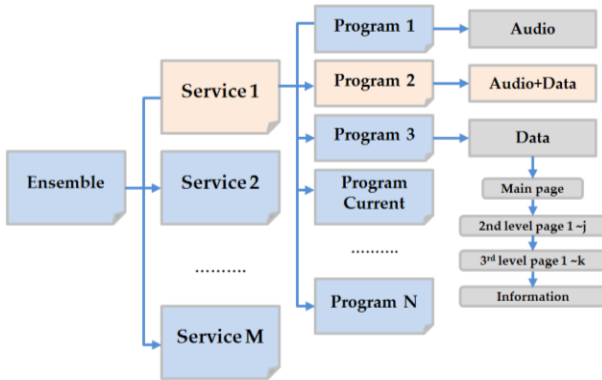


그림 1. DAB 서비스 구조

그림 1과 같이 앙상블 내에는 하나 이상의 서비스들이 다중화되며 해당 서비스들은 시간대별로 하나 이상의 프로그램들로 구성된다. 또한 각 프로그램들은 특정 노래와 같이 프로그램 아이템들로 구분되며 각 프로그램들은 기본적으로 오디오 단독, 오디오 + 데이터 및 데이터 단독 형태로 구성되어 제공된다. 데이터 서비스의 경우에도 그림과 같이 주 페이지(정보)를 시작으로 하나 이상의 하부 페이지를 가질 수 있다. 이러한 구조를 가지고 제공되는 있는 항목(item)들의 접근하기 위해서는 여러 단계의 선택 과정을 통해 이루어진다. 디스플레이와 1 차원 방향 키(navigation) 및 회전(rotary control)을 통해 선택할 수 있는 단말의 기능을 고려할 경우, 청취자는 화면을 보면서 조정 키를 같이 조정하면서 해당 항목에 접근할 수 있다. 이러한 과정은 운전자의 주의를 분산시켜 예기치 않는 상황을 유발할 수 있다. 운전 중 DMB 사용을 허가하지 않은 이유 중 하나이다.

디지털 라디오 서비스의 경우, 서비스 수신 환경을 고려하면 청취자는 주로 청각에 의지하다 필요에 따라 화면을 잠시 주시하는 형태로 그려볼 수 있다. 이를 감안하여 비주얼 라디오[2]가 DMB 규격으로 제정되어 서비스되고 있으며 슬라이드쇼(slide show)를 이용하여 동일한 효과를 제공할 수 있다. 하지만 라디오의 경우, 청취자는 주로 프로그램 아이템과 관련된 부가 정보 즉 프로그램 명, 가수 명, 곡명, 곡에 대한 부가 정보 등이 더 유익한 정보가 될 수 있다. 이러한 정보에 쉽게 접근하고 수신할 수 있는 환경을 제공한다면 사용자는 편리하면서도 안전하게 디지털 라디오를 청취할 수 있다.

이를 위해서는 기존의 비주얼 방식의 사용자 상호작용에 부가적으로 음성 및 터치와 같은 상호작용이 가능하면 직관적인 서비스 제공이 가능하다. 일례로 음성의 경우, 그림 1에 예시된 구조에 대해 음성 키워드를 예약하고 이를 음성 인식 기능을 통해 특정 항목에 접근하도록 한다면 화면을 시청하거나 버튼이나 키를 조작하지 않고도 해당 항목에 접근하여 수신할 수 있다. 또한 음성과 버튼이 동시에 가능할

경우, 음성 키워드를 말하면서 동시에 선택 버튼을 눌러 쉽고 빠르게 해당 아이템을 선택할 수 있다.

3. 모바일 환경에서의 사용자 상호작용

3.1. 멀티모달 상호작용 구조

멀티모달 상호작용은 개념적으로 특정 인스턴스에 하나 이상의 입.출력 인터페이스를 이용하여 단말상에서 제공되는 항목에 접근하여 선택하는 작용을 말한다. 구체적으로는 음성으로 명령을 하면서 동시에 버튼을 누르거나 또는 음성으로 명령을 하면서 화면을 터치 하는 형태의 상호작용을 의미한다. 이러한 기능은 단말 자체에서 수신된 프로그램 정보를 이용하여 고유의 방법으로 구현이 가능하지만 흐름 제어에 있어 매끄럽지 못하거나 때론 불완전한 실행(execution) 모드로 들어갈 가능성이 크며 특히 음성으로 안내할 문장 구성의 경우 자연스럽지 못한 형태로 출력될 수 있다. 이러한 점을 감안하면 EPG 정보를 XML 스키마로 정의하여 제공하는 것과 같이 방송사에서 또는 서비스 제공자가 직접 제작하여 제공하는 것이 효율적이다.

멀티모달 상호작용을 해결하기 위한 방법 중 대표적인 표준 기술로 VoiceXML[3]이 있다. 이는 사용자와 단말 시스템간의 상호작용을 대화(dialog)라는 엘리먼트로 정의하고 연속적이면서 전체적이고 유기적인 대화의 흐름을 하나 이상의 대화로 연결해 놓은 기술이다. 즉, 하나의 대화에서 다른 대화로 이동, 하나 이상의 대화 모음인 페이지에서 다른 페이지로의 이동, 특정 입력에 따른 지정된 페이지 또는 대화로의 이동 등 마치 인간과 인간의 자연스러운 대화와 각 대화마다 정보를 제공하거나 특정 동작(action)을 수행하는 기능을 정의하고 있다. 그림 2는 특정 대화에서 다른 대화로 또는 특정 문서로의 이동(Transition)에 전술한 예를 개념화한 그림이다.

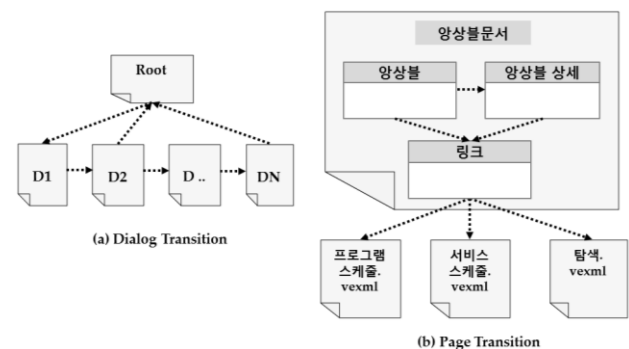


그림 2. 실행 흐름 개념도

기본적으로 루트 문서를 가지며 이는 시작 문서로서 자체적으로 지정할 수 있다. 각 문서는 적어도 하나 이상의 대화(다이얼로그)로 구성되며 문서 내에서 다른 다이얼로그로의 이동이나 다른 문서내의 다이얼로그로 이동이 가능하다. 또한 문서에서 다른 문서로의 이동이 가능하다. 그림 2의 (a)는 다이얼로그에서 다이얼로그로의 이동에 대한 예이며 (b)는 문서에서 문서로의 이동을 도식화 한 것이다. 맨 처음 앙상블 문서가 시작되면 사용자는 앙상블에 대한 상세 정보를 원할 경우 상세 문서로 이동이 가능하며 프로그램 정보를 원할 경우 앙상블 내의 프로그램을 계층적으로 표현하고 있는

문서로 이동할 수 있다. 이때 미리 예약된 음성 키워드 및 특정 버튼을 이용하면 쉽고 빠르게 해당 문서로 이동이 가능하다. 또한 전역 다이얼로그를 지정함으로써 어느 문서에서나 해당 음성 명령을 입력하면 지정된 문서나 해당 다이얼로그로 바로 이동할 수 있는 기능을 제공한다. 이러한 기능을 DAB 즉 디지털 라디오에 적용을 할 경우, 간편하게 원하는 서비스나 프로그램에 접근 할 수 있다.

그림 3 은 대화 엘리먼트 (다이얼로그)에 대한 XML 예문에 대한 예를 보여주고 있다.

```
<?xml version="1.0" encoding="euc-kr"?>
<xml version="1.0" application="vxml.vxml">
  <!-- 음성용 다이얼로그 -->
  <dialog id="mainDialog">
    <grammar type="application/jsgf">상세</grammar>
    <start type="application/jsgf"> 1 </start>
    <prompt timeout="5s">
      라디오 One 앙상블입니다. 원 앙상블에는 4채널의 서비스가 있습니다. 원
      앙상블에 대한 상세 정보를 원하시면 상세라고 말씀하시거나 1번을 누르
     십시오.
    </prompt>
    <block>
      <goto next="specific"/></block>
    </dialog>
    <!-- 음성용 상세 다이얼로그 -->
    <dialog id="specific">
      <block>
        앙상블 상세 정보입니다. 원 앙상블은 라디오 One 방송국에서 제공하는 앙상블
       이며 기본 주파수는 201.072MHz입니다. 프로그램, 장르 및 서비스 제공과
        관련한 문의를 원하시면 www.mbc.co.kr로 접속을 하시거나 042-866-
        1310으로 전화를 주십시오.
      </block>
      <goto next="link"/>
    </block>
  </dialog>
</xml>
```

그림 3. 다이얼로그 예문

일반적인 다이얼로그는 <form> 엘리먼트를 사용하며 그림과 같이 <form> ID 를 가지고 식별을 한다. 입,출력을 위한 <field>를 하부 엘리먼트로 가지며 입력 명령을 규정하고 있는 grammar 및 문자를 음성으로 출력하기 위한 <prompt> 엘리먼트로 구성된다. 그림 3 의 예에서는 Radio One 앙상블에 대한 설명을 음성으로 제공하고 청취자가 상세라고 말하거나 1 번을 누르면 그림 2 의 흐름과 같이 해당 상세 정보로 이동하는 것을 보여주고 있다. 다음 절차로는 <goto> 엘리먼트에 지정된 다이얼로그인 "link"로 이동을 한 후 프로그램 스케줄에 대한 정보를 제공해 주는 흐름으로 구성되어 있다.

3.2. 디지털 라디오 환경에서의 멀티모달 상호작용

3.1 에서 기술한 XML 기반의 멀티모달 상호작용 기술을 방송환경에 적용하기 위해서는 다음과 같은 고려사항이 필요하다. 어플리케이션 생성, 어플리케이션 구성, 음성 입,출력, 음성 명령어 및 전송 등이 기술적인 고려사항에 해당한다.

어플리케이션 생성은 VoiceXML 언어를 사용하여 “.vxml” 확장자를 가진 문서를 생성함으로써 가능하다. 이때 고려할 사항으로는 기존 VoiceXML 규격과 호환성을 보장할 경우, 정의된 이름공간(namespace) 및 스키마를 그대로 이용하면 되지만 만약 새로운 엘리먼트(element)나 속성(attribute)을 지정할 경우, 이름공간 및 스키마의 변경이 요구된다. 어플리케이션 생성은 디지털 라디오 환경의 전송 용량을 고려하여 가능하면 간략하게 생성하는 것이 바람직하다. 이를 위해 각 엘리먼트 및 속성을 바이너리 EPG 와 같이 바이너리화 하면 압축 효과가 있어 효율적이다.

다음으로 어플리케이션의 구성을 살펴보면, 디지털 라디오는 하나의 앙상블 대역 내에 하나 이상의 서비스 들이 다중화 되어 있으며 해당 앙상블 내에 이 어플리케이션을 전송할 경우

다른 앙상블에 대한 정보가 없어 전체 앙상블 서비스를 제공하는데 제약이 있다. 이를 위해서는 단말에서 자체적으로 앙상블을 연결해주는 최 상위 문서를 구성한 후 각 앙상블의 루트 문서와 링크를 명시해 주면 통합 서비스를 제공할 수 있다.

음성 입,출력의 경우, VoiceXML 에서 지원하는 ASR(Automatic Speech Recognition), TTS(Text To Speech) 및 음원 파일 재생 기능을 이용할 수 있다. 이와 더불어 간단한 숫자를 지정하여 선택할 수 있는 버튼 입력이 가능하다.

음성 명령어는 매우 유익한 기능으로 VoiceXML 에서는 제한을 두지 않고 있다. 심지어 문장에 대한 해석을 통해 키워드를 도출한 후 이에 반응하는 기능을 지원하고 있으나 이러한 인식 기능은 단말 자체에서의 성능 한계로 전용 인식 서버를 운용하는 원격 시나리오에 적합하다. 방송 환경의 단말에는 비교적 인식이 용이한 키워드를 사용하는 것이 적합하다. 이는 차량 환경에서 발생하는 잡음을 고려하면 이에 대한 실제 테스트를 통해 선정하는 것이 바람직하다.

또한 음성 명령의 경우, 미리 예약된 키워드를 사용하고 그에 대한 전역 천이(Global Transition)를 미리 지정할 경우 복잡한 단계를 거치지 않고 직접 해당 정보를 담고 있는 문서나 다이얼로그로 이동할 수 있는 장점을 제공한다. 이를 위해서 디지털 라디오에서 주로 사용되고 있는 음성 키워드를 예약하여 사용하면 상호작용에 대한 성능을 극대화 할 수 있다. 일 예로 “앙상블”을 키워드로 지정하고 청취자가 어느 문서에서나 “앙상블”을 입력하면 지정된 앙상블 문서로 바로 이동 할 수 있다. 이는 청취자가 단말 시스템과 대화를 진행하다 앙상블에 대한 정보로 돌아가고 싶을 때 여러 단계를 통해 복귀할 필요 없이 바로 이동할 수 있는 장점을 제공한다.

전송의 경우, 디지털 라디오의 특성을 감안하면 다음과 같은 사항을 고려해 볼 수 있다. 전송 효율을 고려하여 어플리케이션을 구성하는 모든 파일을 zip 형식으로 묶어 전송하면 발생 비트 율을 낮출 수 있다. 실제 전송은 MOT(Multimedia Object Transfer) 규격을 적용하여 단일 데이터 서비스로 제공한다. 만약, 브로드밴드 네트워크 가능할 경우에는 특정 문서들은 브로드밴드 망을 이용하여 부가적으로 수신 받을 수 있는 형태로 전송하는 것도 효율적이다. 기술적으로는 <goto>나 <link>와 같은 transition 엘리먼트 속성에 해당 URL(Uniform Resource Locator)을 지정하면 별도의 방법에 대한 기술 없이 단말 시스템에서 이에 대한 기능을 구현함으로써 가능하다.

3.3. 시스템 구조

기술한 멀티모달 상호작용 어플리케이션은 VoiceXML 엔진과 하부에 존재하는 컴포넌트인 ASR, TTS, Button Interface 및 음원 파일 재생 기능에 대한 기능의 구현으로 어플리케이션 구동이 가능하다. VoiceXML 엔진은 해석기가 핵심 기술로 정의된 엘리먼트를 해석하고 다이얼로그에 대한 전체 흐름을 제어하는 기능을 담당한다. 또한 해석기는 항상 사용자로부터의 입력을 대기하며 입력 이벤트가 발생하면 이를 체크한 후 각 다이얼로그에 지정된 입력 명령어와 비교한 후 일치된 다이얼로그를 실행시킨다. 특정 음원 파일에 대한 재생이 필요하거나 음성을 변환해야 할 문장이 있으면 하부 컴포넌트를 통해 해당 문장을 음성을 출력하도록 이를 제어

한다. 그림 4는 이에 대한 개념도를 도시하고 있다.

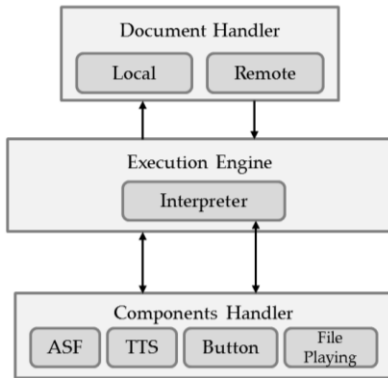


그림 4. 실행 엔진 구조도

3.4. 어플리케이션

적용 가능한 어플리케이션은 기존 비주얼 방식의 어플리케이션이 제공할 수 있는 많은 범주의 정보를 제공할 수 있다. 그림 5와 같이 기본적으로 EPG를 음성으로 안내할 수 있을 뿐만 아니라, 디지털 라디오에 핵심 서비스라고 할 수 있는 교통정보, 뉴스, 일기 예보, 교육, 지역 정보 및 음성 게임(voice games) 등이 가능하다.

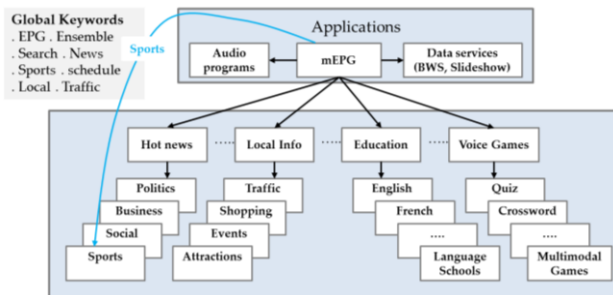


그림 5. 적용 가능 어플리케이션

mEPG는 멀티모달 EPG를 의미하며, 시작 문서로 지정되어 있다. 프로그램 안내 문서에서 시작하여 방송되는 오디오 채널 및 관련 데이터 서비스에 접근할 수 있으며 자체적으로 구성된 정보 문서로 이동하여 관련 정보를 들을 수 있다. 또한 예시한 바와 같이 어느 문서에 위치해 있던 지 “global keyword”에 대한 음성 입력을 할 경우, 바로 해당 문서로 이동하여 해당 정보를 들을 수 있는 장점을 제공한다. 그림 5에서는 청취자가 어느 곳에서나 “sports”란 키워드를 입력 하면 바로 “sports” 정보를 담은 문서로 이동하여 첫 번째 다이얼로그부터 단말 시스템과의 상호작용을 시작할 수 있다.

3.4 구현 및 시험

기능 검증을 위해 PC 환경에서 open source로 제공되는 VoiceXML SDK를 이용 하였으며 TTS 및 ASR은 상용 솔루션을 적용하였다. 어플리케이션에 대한 실행은 MOT 전송을 미리 수신된 것으로 가정하여 로컬 환경으로 구동하였다. 그림 6은 앙상블 문서를 화면으로 출력한 예로

청취자가 “Radio One”을 명령하면 Radio One 프로그램의 상세한 정보를 포함하고 있는 문서로 이동을 하여 해당 정보를 음성으로 제공한다. 그림에서 음성 키워드는 “Radio One”, “Radio Two”, “program” 및 “search”로 만약 “search”를 입력할 경우 프로그램 검색 문서로 이동하여 청취자가 원하는 프로그램을 검색할 수 있다. 버튼 입력으로 “1”번을 누르면 “Radio One” 문서로 바로 이동을 한다. 또한 어느 페이지에서든지 “앙상블” 키워드를 명령하면 “앙상블” 문서로 바로 이동할 수 있다.

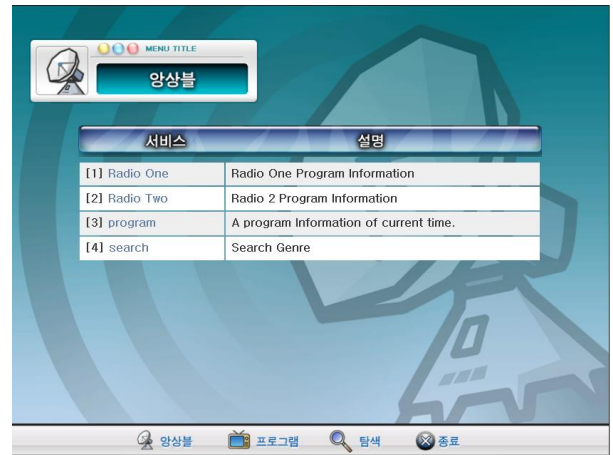


그림 6. 앙상블 문서 화면

4. 결론

본 논문에서는 디지털 라디오와 같은 이동 환경에서 기존 비주얼 기반의 사용자 상호작용만이 아닌 음성 기반의 상호작용을 포함한 다양한 입.출력 상호작용이 가능한 멀티모달 기반의 상호작용에 대해서 연구를 수행하였다. 먼저 모바일 방송 환경을 분석하여 이에 적합한 요구사항을 도출하였으며 이를 기반으로 적합한 사용자 상호작용 시스템 구조를 제안하였다. 제안한 사용자 상호작용 기술을 적용할 경우, EPG뿐만 아니라 디지털 라디오 방송에서 제공 가능한 다양한 정보를 쉽고 편리하게 제공 가능함을 확인하였다.

감사의 글

본 연구는 미래창조과학부 및 정보통신기술진흥센터의 정보통신·방송 연구개발 사업의 일환으로 수행하였음. [2014-044, USAC 기반의 디지털라디오 송수신 시스템 기술 개발]

참조문헌

- [1] ETSI EN 300 401 V.1.4.1 (2006-1) Radio Broadcasting Systems; digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers
- [2] ETSI ES 102 428: “DAB; DMB video service; User application specification”, April 2009
- [3] Voice extension markup language (voicexml) version 2.0. W3C Recommendation, 16 March 2004. <http://www.w3.org/TR/voicexml20/>.