

TTS를 이용한 매장 음악 방송 서비스 시스템 구현*

Implementation of Music Broadcasting Service System in the Shopping Center Using Text-To-Speech Technology

장 문 수** · 강 선 미***
Moonsoo Chang · Sunmee Kang

ABSTRACT

This thesis describes the development of a service system for small-sized shops which support not only music broadcasting, but editing and generating voice announcement using the TTS(Text-To-Speech) technology. The system has been developed based on web environments with an easy access whenever and wherever it is needed. The system is able to control the sound using silverlight media player based on the ASP .NET 2.0 technology without any additional application software. Use of the Ajax control allows for multiple users to get the maximum load when needed. TTS is built in the server side so that the service can be provided without user's computer. Due to convenience and usefulness of the system, the business sector can provide better service to many shops. Further additional functions such as statistical analysis will undoubtedly help shop management provide desirable services.

Keywords: music broadcasting, TTS(Text-to-Speech), web based system, voice announcement

1. 서 론

문명이 발달하여 사람들의 생활수준이 높아짐에 따라 소비에 대한 개념이 단순히 필요에 의한 것에서부터 심리적 충족을 위한 것으로 변화되었다. 그러므로 소비자들의 마음을 사로잡기 위한 인터넷 마케팅, 아로마 마케팅, 음악 마케팅 등 다양한 마케팅 전략들이 중요하게 대두되고 있다. 그 중 음악 마케팅은 소비자의 감성을 촉구하는 감성 마케팅의 대표적인 것으로서 소비 촉구 및 소비자와 판매자 간의 친밀도 향상 등에 큰 영향을 미쳐 크고 작은 규모의 매장에서 없어서는 안될 중요한 요소로 자리잡고 있다.

최근 음악 방송에서 음원에 대한 저작권 침해에 관한 논란이 발생되면서 음반 판매 매장, 마트, 공원, 영화관 등의 중소규모 매장에서 음원 사용료를 지불하지 않은 음악 방송 행위를 불법으로 정

* 본 논문은 중소기업 산학협력실 지원사업의 연구결과로 수행되었습니다.

** 서경대학교 소프트웨어학과

*** 서경대학교 컴퓨터학과, 교신저자

의하였다. 대부분의 중소규모 매장은 음원 사용료를 지불하기에는 부담이 되지만, 중요한 마케팅 방법의 하나인 음악 방송을 중단할 수 없으므로 이에 대한 적절한 해결방안이 연구되고 있다. 이에 일정한 회비를 받고 음원 사용료를 대신 지불한 후 각 회원 매장에 음원을 제공하는 서비스를 하는 업체들이 생겨나기 시작했다. KT의 Shopcast(shopcast.bizmeka.com), SK Telecom의 Biz melon(bizmelon.com), Biz music(bizmusic.co.kr), 티온(teeon.net) 등이 그 대표적인 서비스들이다.

이러한 음악 방송 서비스 시스템은 음원 사용에 대한 해결책을 제시하여 기존 매장에서 각자 보유하고 있던 음원을 재생해주던 것과 동일한 역할을 지원하고 있다. 그러나, 매장 측에서는 기존에 무료로 사용했던 매장 음악 방송 서비스를 유료화한 것이므로, 이러한 비용에 대한 보상으로 새로운 기능을 서비스에 추가할 것을 요구하고 있다. 가장 많이 알려진 Shopcast에서는 음악방송을 위한 다양한 서비스 외에도 셋탑박스를 이용한 서비스도 지원되고 있으며, Biz melon에서는 모바일을 연동하여 매장을 방문한 고객이 문자로 해당 매장에 음악을 신청할 수 있는 서비스를 제공하고 있다. 이와 같은 추세는 가입회원 업체들의 보상 요구에 대한 응답이라고 볼 수 있다.

본 논문은 기존의 매장 음성 방송 서비스에 안내방송 지원을 위해 음성합성기(TTS: Text-To-Speech) 기술을 접목한 서비스 지원 방식과 사용자가 쉽게 이용할 수 있는 웹 환경에서의 시스템 구현에 관한 것이다. 사용자 측에서는 개점 방송이나 폐점 방송 등 원하는 방송 문구를 서버에 입력하면, 지정 시간에 요구한 안내방송을 서버측 데이터베이스에 저장된 음원에 추가하여 서비스를 요구한 매장에서 자동으로 방송된다. 또한, 매장에서 행해지는 각종 행사에 적합한 안내방송들을 사용자가 수시로 편집할 수 있고, 전문가에 의해서 제작된 특별한 안내방송을 서버에서 제공하여 사용자가 이용할 수 있게 한다. 이와 같은 서비스는 매장에서 직접 마이크로 실시간 방송이나 녹음 방송을 해왔던 것을 대체하여 간편한 조작만으로 서비스가 가능하므로 사용자 측에서의 편의를 최대한 고려한 것이며, 음원 제공 서비스업체 측에서도 매장에서 요구하는 방송을 쉽게 지원하는 서비스를 제공함으로써 각 매장들로 하여금 회원 가입을 촉진시키는데 기여하게 된다[1][2].

2. 제안하는 매장 음악 방송 서비스

본 논문에서 제안하는 매장 음악 방송 서비스 시스템은 마트나 음반 판매 매장, 공원 등 중·소규모 매장에서 컴퓨터 작업 환경에 익숙하지 않은 운영자를 대상으로 개발한다는 전제 하에, 사용자가 친숙하고 편리하게 요구되는 정보를 입력하도록 웹 환경을 선택한다. 본 서비스 시스템은 서버-클라이언트 시스템으로 구성되어 각 클라이언트는 웹 브라우저를 통하여 서비스를 제공받는 형식으로 구현된다. 사용자들은 인터넷이 연결되어 있는 환경에서는 어디서라도 동일한 서비스를 제공할 수 있으며, 개인에 대한 설정 정보도 서버에서 유지하여 어디서든지 같은 설정 값을 유지한다. 본 장에서는 제안하는 매장 음악방송 서비스에 대하여 서버에서 클라이언트로 제공되는 서비스의 종류에 따라 여러 가지 기능을 조합하여 최적의 서비스 환경을 구성하도록 제안한다.

<그림 1>은 사용자가 웹 인터페이스를 통하여 서버로부터 제공받는 서비스의 전체 흐름도를 나타낸다.

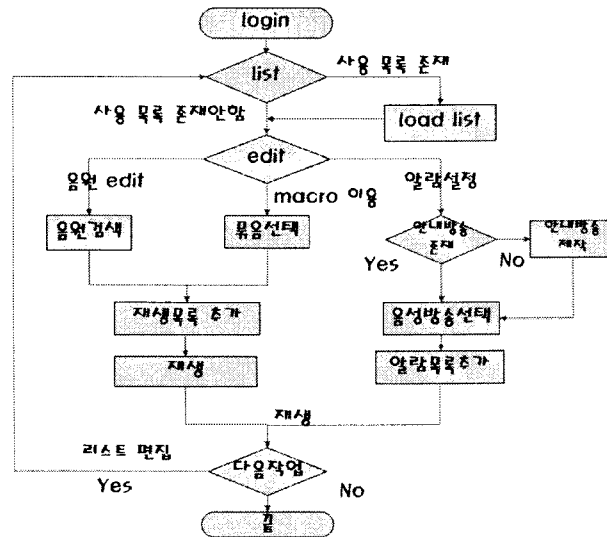


그림 1. 제안하는 시스템의 서비스 흐름도

클라이언트는 웹에 접속하여 회원 확인 후 사용자가 원하는 음원 재생 목록을 작성할 수 있도록 한다. 사용자는 여러 개의 재생 목록을 가질 수 있어서 요일별로 재생 목록을 다르게 설정해 두거나 특정 날짜에 필요한 목록을 미리 만들어 사용할 수 있도록 한다. 음원의 검색은 음악의 경우 제목, 가수명의 일부를 이용하여 검색이 가능하고, 카테고리별로 구분되어 있어서 시대별이나 장르별 검색도 가능하게 한다.

안내방송의 경우 클라이언트가 미리 작성해 놓은 안내방송들 중 선택하여 사용할 수 있도록 한다. 안내방송은 알람목록으로 추가되는 것을 전제로 하며 세 가지 형태가 있다. 첫째는 배경음악 없이 안내방송만 나오도록 한 것이고, 둘째는 재생되고 있는 음악을 배경음악으로 안내방송이 재생되는 형태로서 재생 음악 볼륨을 적당한 수치로 줄여 안내방송과 어울려 방송될 수 있도록 한다. 셋째는 지정한 음악을 배경 음악으로 사용하여 안내방송을 하는 경우이다. 배경 음악이 흘러나오는 중 언제 안내방송을 서비스할 것인가를 지정하게 한다. 이 서비스의 경우는 세 개의 미디어플레이어가 동시에 사용되며, 음악의 재생, 음성(안내방송)의 재생, 배경음악의 재생으로 이후의 설명을 위해 각 미디어 플레이어들을 음악 플레이어, 음성 플레이어, 배경음악 플레이어라 명명한다.

또한 클라이언트는 관리자가 제공하는 음악과 안내방송을 사용할 수 있다. 이것은 현재 매장 음악 방송 서비스를 하고 있는 많은 업체에서 제공하는 서비스의 형태로 일반적으로 채널 혹은 테마라고도 부르지만, 여기서는 좀 더 직관적으로 알 수 있도록 노래목록이라는 용어를 사용한다. 음원의 재생 리스트는 하나의 데이터베이스의 재생 목록 테이블에 저장되지만, 사용자 정보를 이용하여 각 클라이언트는 해당 클라이언트의 재생 목록만을 보게 한다.

2.1 알람 기능 서비스

알람은 특정 시간에 원하는 안내방송을 할 수 있는 기능을 말한다. 미리 생성한 안내방송은 간단한 조작으로 알람 지정이 가능하도록 설계된다. 알람은 서비스 형태에 따라 세 가지로 방송되는데,

첫째는 가장 단순한 형태인 안내방송만이 재생되는 것이고, 둘째는 음악 플레이어를 통해 재생되는 음원을 배경음악으로 하여 안내방송이 되는 형태이고, 셋째는 지정 배경음악을 갖는 형태이다. 여기서 지정 배경음악은 “광파레”나 “축하음”과 같이 특수 효과를 줄 수 있는 음원이나 사용자가 지정할 음악을 의미한다.

첫 번째 형태의 구현 방법은 알람 시간이 되면 음악 플레이어의 볼륨을 서서히 줄여 fade out의 효과를 낸다. 볼륨이 0이 되면 안내방송 플레이어가 실행되고, 알람의 재생이 끝나면 다시 음악 플레이어의 볼륨을 서서히 높여 fade in의 효과를 낸다. 두 번째 형태는 음악 플레이어의 볼륨을 서서히 줄여 fade out의 효과를 내는 것은 첫 번째 형태와 동일하나 볼륨을 음소거의 수준이 아닌 적당한 수준으로 줄인 후 안내방송 플레이어가 실행되게 한다. 세 번째 형태의 구현 방법은 첫 번째와 마찬가지로 음악 플레이어의 볼륨을 0으로 만들어 소리를 없앤 후, 사전에 정해 놓은 지정 배경음악을 이용하여 방송한다. 여기서는 안내방송 플레이어와 배경음악 플레이어가 동시에 재생된다. 알람에 추가된 안내방송은 사용자가 원할 때 제거 및 변경이 가능하다. 알람의 실행 순서도는 <그림 2>와 같다.

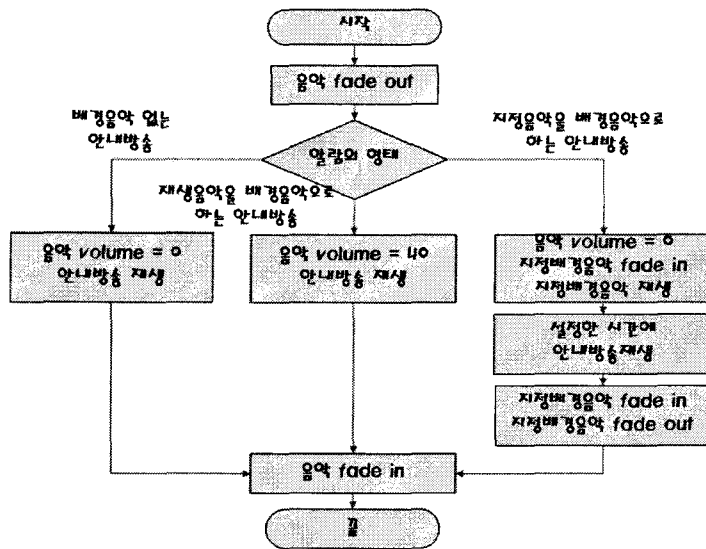


그림 2. 제안하는 시스템의 알람 실행 순서도

2.2 안내방송 지원 서비스

본 시스템의 핵심 서비스인 안내방송 지원은 알람 기능과 함께 재생된다는 전제하에 구현된다. 안내방송은 클라이언트 측에서 원하는 문장을 서버로 전송하면 서버 측에 탑재되어 있는 TTS를 통하여 PCM으로 만들어진 데이터를 mp3 형태로 변환한다.

클라이언트는 해당 문장을 안내방송 음성으로 만든 후 바로 들어 볼 수 있고, 알람목록으로 추가하기 전에 클라이언트가 만든 안내방송들의 목록에 삽입하여 별도로 관리할 수도 있다. 해당 안내방송은 모두 실시간으로 작성, 수정 및 삭제가 가능하여 클라이언트 측에서는 유동성이 있는 안내방송을 통해 만족도 높은 방송을 할 수 있도록 한다. 상기에 언급한 알람의 설정은 이 과정에서 이루어진다.

본 시스템에서는 첫 번째와 두 번째 형태의 알람에 추가되는 안내방송을 단순 안내방송으로, 마지막 지정된 배경음악을 이용한 형태의 알람에 추가되는 안내방송을 지정 안내방송으로 구분한다. 단순 안내방송은 재생되는 음원을 계속 재생하는 것에 대한 제어만 설정하면 되지만, 지정 안내방송의 경우에는 사용자가 지정한 배경음악이 재생되고 있는 도중 원하는 시점에서 안내방송이 시작되어야 한다.

제안하는 시스템에서는 사용자가 지정한 배경음악을 듣는 도중 원하는 시점을 지적하면, 시작 시점부터 지정된 시점까지의 시간을 계산하여 사용자가 원하는 배경음악과 안내방송의 재생 시간이 자동으로 맞춰지도록 한다. 안내방송과 배경음악의 길이가 서로 다를 것으로 예상하여 사용자가 원하는 시점에서 배경음악의 재생을 끝낼 수 있도록 상기와 동일한 방법으로 재생의 끝점을 지정할 수 있도록 한다. 이 기능은 사용자를 위해 제공되는 기능이지만 조작이 간단하지 않다는 문제점이 있으므로 배경음악 시작 5 초 후 안내방송이 시작되고, 안내방송이 완료된 5 초 후 배경음악을 완료하는 자동 모드를 추가한다.

매크로는 관리자가 사용자의 편의를 위해 제공되는 기능으로 특정 노래 묶음과 일반적인 안내방송이 들어있다. 노래 묶음은 일일이 음원을 선택하는 유동적인 방송을 원하지 않는 사용자들을 위해 추가한 것이다. 일반적인 안내방송은 자주 사용되는 안내방송을 미리 만들어 제공함으로써 사용자들이 보다 쉽게 방송을 할 수 있도록 하는 것이다. 노래 묶음은 장르별, 시대별 등으로 제공되며 사용자는 노래 묶음 안에 어떤 음원들이 포함되는지를 상세 검색창을 통해 확인할 수 있다. 노래 묶음이 재생목록에 추가되면 노래 묶음에 포함되어 있는 음원들이 각각의 이름으로 추가된다. 이것은 사용자가 재생목록에 포함된 노래 묶음 안의 내용을 기호에 맞게 편집할 수 있도록 하기 위함이다.

음원 컨트롤은 일반적인 미디어 플레이어와 동일한 기능을 갖도록 구현한다. 재생, 일시정지, 정지 및 한곡 건너뛰기, 한곡 앞으로 가기의 기능이 있으며, 볼륨은 수시로 조절할 수 있도록 한다.

3. 시스템 구성 제안

<그림 3>은 제안하는 매장 음성 방송 시스템의 구성도로 서버와 클라이언트가 각각 지원하는 서비스를 표시한다.

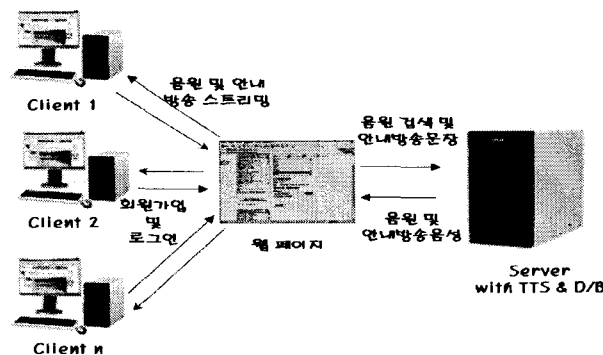


그림 3. 제안하는 매장 음악 방송 서버와 클라이언트 구성도

3.1 시스템 버전과 개발 도구 및 라이브러리

본 시스템의 개발도구는 Visual studio 2005이며, 데이터베이스는 MS SQL 2005[3][4], 서버는 Windows 2003 server[5]를 사용한다. 웹 페이지 형태로 구성되며, Microsoft사의 최신 기술인 ASP .NET 2.0[6][7]을 이용한다. 서버 컨트롤의 사용으로 인하여 시스템의 부하가 커지는 부분은 Ajax[8][9][10] 컨트롤을 클라이언트 컨트롤로 대체함으로써 시스템을 최대한 가볍게 설계하여 이용자들이 증가할 경우를 대비하여 예상되는 부하를 최대한 고려한다. TTS는 서버용으로 제작된 (주)미디어젠 음성합성기 2.1 2GB 용량[1]을 장착하고, 미디어 플레이어는 MS사의 SilverLight alpha 1.1[1]을 사용하여 구현한다. 합성기에서 출력되는 row PCM 파일(16kHz/mono)을 mp3 파일로 변환하는 프로그램은 Lame 인코더의 라이브러리[11][12][13]를 사용한다.

3.2 제안하는 시스템의 데이터베이스 구성

데이터베이스 테이블은 사용자들을 관리하는 member, 재생 목록을 관리하는 song schedule, 알람 목록을 관리하는 alarm schedule, 매크로를 관리하는 macro, 알람의 배경음악을 관리하는 voice_BGM, 음원을 관리하는 테이블인 song, 안내방송 음원을 관리하는 voice가 있다. <그림 4>는 제안하는 시스템의 데이터베이스 구성도를 나타낸다.

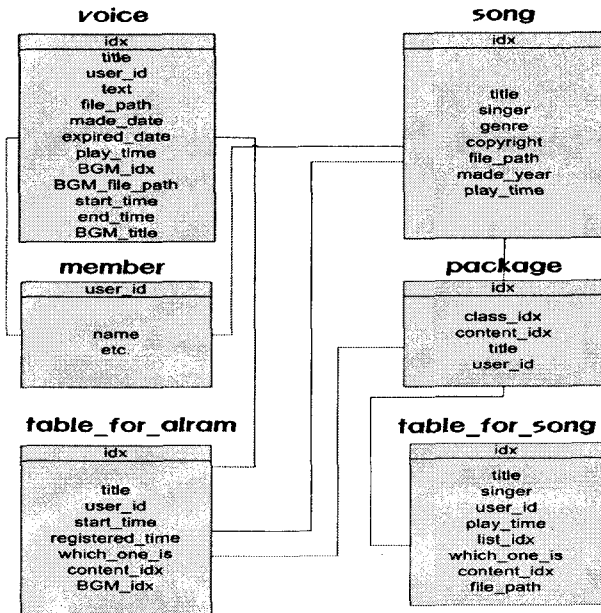


그림 4. 제안하는 시스템의 데이터베이스 구성도

TTS로 생성되는 안내방송 음원을 관리하기 위한 voice는 합성문장의 생성 날짜(made_date)와 만료 날짜(expired_date)를 자동으로 지정한다. 만료 날짜는 최초 생성 날짜를 기준으로 1주일 후

1) Microsoft silverlight(기존 코드명 "WPF/E")는 웹에서 사용할 차세대 .NET 기반 미디어 환경 및 풍부한 대화형 애플리케이션을 제공하기 위한 다중 브라우저 다중 플랫폼 플러그인이다.

로 지정되고, 이후 안내방송이 사용되면 만료 날짜가 사용 날짜로부터 1주일 후로 자동으로 수정된다. 만료 날짜가 된 안내방송 음원들은 사용자에게 삭제 여부를 확인하게 하여 재사용하지 않을 음원들은 삭제되도록 한다. 이것은 자주 사용되는 안내방송은 DB에 저장하여 정보를 유지하고 사용되지 않는 것은 삭제하여 서버의 부담을 덜기 위함이다. 각 사용자별로 작성한 안내방송은 리스트로 확인, 추가 및 삭제가 가능하다. 또한, macro는 관리자에 의해 제공되는 서비스들의 묶음으로 매일 서비스 전문 관리자에 의해 갱신되는 노래묶음이나 자주 사용되는 인사말이나 오늘의 날씨 안내 등을 제공하기 위한 것이다. 그 외에 alarm_schedule은 매장에서 특정 시간에 특정 안내방송이나 행사성 음악 방송을 원할 경우, 시간 정보를 저장하기 위한 것이다.

3.3 음성 합성기 처리 모듈

시스템에 사용된 음성 합성기는 입력된 문장을 자연어처리에 의해 기본 음성단위로 분할하여 버퍼에 저장한 후, 8,000 개의 문장으로 구성된 음성화일 데이터베이스에서 각 음성단위에 가장 유사한 PCM 데이터를 가져와서 순서대로 연결하여 음성화일을 구성한다. 본 시스템에서는 스트리밍 서비스를 위해 선택된 PCM 데이터를 mp3 파일 형태로 변환하여 저장한다. 합성기 처리 속도를 고려하여 입력 문자열의 길이는 512 자를 넘지 않도록 구성되어 있고, 만약 문장의 길이가 512 자를 넘을 경우에는 다수개의 블록으로 나누어서 만들어진 mp3 파일을 버퍼에 저장했다가 최종 안내방송 파일을 만드는 방법으로 구현한다.

구현된 합성기 모듈은 사용자가 안내방송 생성을 요구 할 때 초기화 모듈을 구동시키면서 사용자별로 관리된다. 사용자가 요구하는 안내방송들은 사용자 아이디를 이용하여 파일로 생성되고 관리되므로, 동시에 다수의 사용자가 안내방송 합성을 요구할 경우에도 혼돈은 일어나지 않는다. 초기화 작업이 마쳐진 후 합성음 생성 모듈을 실행하여 PCM 데이터를 생성하고 데이터 포인터를 반환하면, 반환된 포인터를 가지고 Lame 인코더를 실행하여 mp3 파일을 생성한다. 합성기는 대기 상태에 있기 때문에 사용자가 원하는 때마다 초기화를 하지 않고 합성음을 생성할 수 있다. 합성기 구동 및 제어 흐름도는 <그림 5>에 나타내었다.

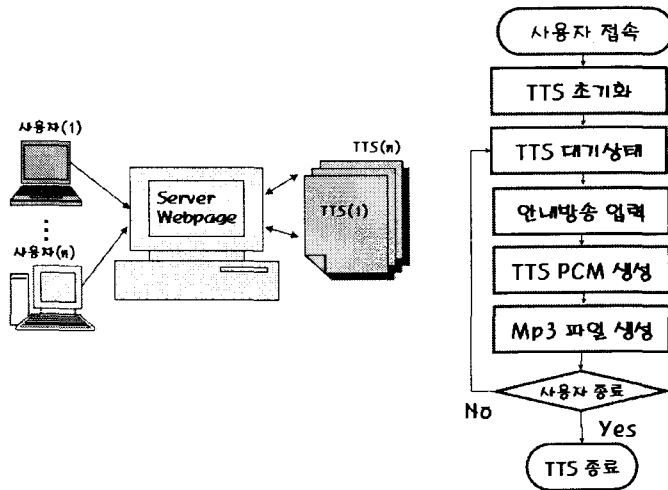


그림 5. 다중 사용자 접속기 TTS 구동 및 제어 흐름도

3.4 웹 인터페이스 구현

클라이언트에는 총 세 개의 미디어 플레이어가 사용된다. 하나는 일반 음원을 시간의 개념 없이 순차적으로 재생하는 플레이어이고, 다른 하나는 알람 시간이 되었을 때 알람을 울리기 위한 플레이어이다. 또 다른 하나는 배경음악을 갖고 있는 지정 안내방송의 배경음을 재생해 주기 위한 플레이어이다. 지정 안내방송의 배경음악 스트리밍을 실제 소리를 포함시켜 하나의 스트리밍으로 제공하지 않고 따로 스트리밍하며 섞어 놓은 듯한 효과만을 준 것으로 방송을 하는 도중에 실시간으로 안내방송의 볼륨과 배경음악의 볼륨을 임의로 조절할 수 있도록 하기 위해서이다.

클라이언트에게 제공되는 서비스인 음원의 재생은 Microsoft사의 신기술인 silverlight 미디어 플레이어를 이용한다. 별도의 어플리케이션 프로그램의 설치 없이 음원을 제어할 수 있으며, 이러한 음원은 서버에서 클라이언트로 mp3 파일 형태로 스트리밍함으로써 네트워크의 트래픽을 줄이도록 설계한다. 알람에서 요구하는 음악의 볼륨 조절과 안내방송 플레이어도 mp3 파일에 대하여 별도의 처리과정 없이 시스템에 있는 컨트롤러의 조합만으로 실행할 수 있다. 또한 fade in, fade out 효과를 통해 방송이 중단되었다는 느낌을 주지 않고 매끄러운 진행이 가능하도록 한다. 구현되는 웹 페이지의 인터페이스는 <그림 6>과 같다.

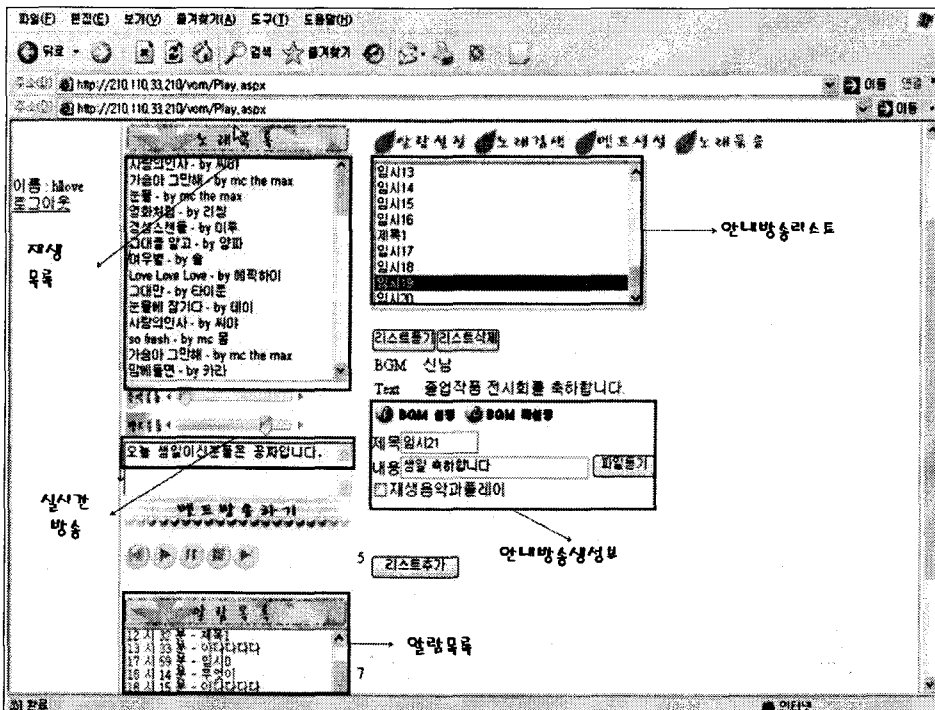


그림 7. 구현되는 음악 방송 서비스 시스템의 웹 인터페이스

4. 결 론

본 논문에서 제안 구현되는 시스템은 음악 방송지원뿐만 아니라, 사용자가 요구하는 각종 안내 방송을 원하는 형태로 만들 수 있으므로 전문 방송 서비스의 효과를 매장에 제공할 수 있도록 한다. 본 시스템에서 구현된 안내방송 제작 서비스는 기존의 매장 음악 방송 시스템과의 결합이 용이하고, 음성 합성기 기술을 응용하여 서비스의 질을 높이는 융합 기술의 좋은 예가 된다. 이러한 시스템은 현재 많은 사랑을 받는 웹 음악 방송에 응용될 수 있다. 메신저에 올려 채팅을 하는 동안 함께 음악을 듣고, TTS를 이용하여 DJ처럼 안내방송을 하는 등의 다양한 서비스의 개발도 가능하다. 또한, 전시회 등에서도 해당 전시 작품에 대한 내용을 소개해주고 관련 배경음악을 제공하는 서비스에도 응용이 가능하다.

향후 저작권 보호를 위한 방법으로 음원 제공자들이 요구하는 각종 광고를 음원에 첨부하여 방송하는 방법으로 음원 서비스의 무료 회원제로의 전환도 가능하도록 한다. 추후 추가되어야 할 기능으로는 서버에서 매장을 관리할 때에 필요로 하는 각종 통계적인 자료를 쉽게 얻을 수 있도록 하는 기능을 들 수 있다.

참 고 문 헌

- [1] 미디어젠. 2006. 음성 HMI 전용 단말기용VUI 모듈 개발 보고서.
- [2] 정경채, 강선미, 장문수. 2004. "지역 서버 기능을 강화한 다중 서버 음성 수집 시스템의 개선 : CSDBWeb3.1," 한국통신학회 하계학술대회 논문집. 47.
- [3] 우재남. 2007. 뇌를 자극하는 SQL SERVER 2005. 한빛미디어.
- [4] Robert Vieira. 2006. *Beginning SQL Server 2005 Programming*. 정보문화사.
- [5] Mark Minasi. 2005. *Mastering windows Server 2003*. 정보문화사.
- [6] 이시환. 2007. 뇌를 자극하는 ASP. NET 2.0. 한빛미디어.
- [7] ASP .NET Forum(forums.asp.net), ASP .NET 2.0.
- [8] Laurence Moroney. 2006. 마이크로 소프트웨어 ASP.NET AJAX. 도서출판 ITC.
- [9] 이 혁. 2007. Semantic web 구현을 위한 Ajax. FREELEC.
- [10] Microsoft ASP .NET Ajax Forums(forums.asp.net), ASP .NET Ajax.
- [11] LAME (Lame Aint an MP3 Encoder), *Open Source Encoder*. <http://lame.sourceforge.net/>.
- [12] 김용성. 2003. VISUAL C++6 완벽가이드. 영진출판사.
- [13] Devpia Visual C++ 포럼 마을(www.devpia.com/VC++.MAEUL), Visual C++ 6.0.

접수일자: 2007. 10. 8

게재결정: 2007. 11. 26

▲ 장문수

서울특별시 성북구 정릉 4동 (우: 136-704)

서경대학교 소프트웨어학과

Tel: +82-2-940-7509 Fax: +82-2-919-5075

E-mail: cosmos@skuniv.ac.kr

▲ 강선미

서울특별시 성북구 정릉 4동 (우: 136-704)

서경대학교 컴퓨터과학과

Tel: +82-2-940-7291 Fax: +82-2-919-5075

E-mail: smkang@skuniv.ac.kr