

웹서비스를 이용한 음성인식시스템

오지영, 김윤중

oz1115@empal.com, yjkim@hanbat.ac.kr

국립 한밭대학교

대전광역시 유성구 덕명동 산16-1, Tel +82-42-821-1143

Keyword : 음성인식, 웹서비스, 웹서버

인터넷은 사회 여러 분야에 걸쳐 다양한 영향을 주고 정보의 교류나 지역 간의 격차를 해소하는데 지대한 공헌을 하였으며 현재는 우리나라 가정의 67%가 초고속 인터넷을 사용할 만큼 급속도로 발전되었다. 또한 음성인식 기술도 많은 연구가 이루어지고 있다. 음성인식 기술은 어떠한 기능이나 제품 등을 동작시키는 기능으로 이미 문서작성 및 편집, 게임, 기계, 모바일 등 다양한 분야에 응용되고 있다. 이렇듯 인터넷과 음성인식 기술이 발전함에 따라 인터넷과 사용자간의 인터페이스로 음성을 이용하고자 하는 욕구가 증대되고 있다.

이와 같이 음성 인터페이스의 기대가 증가함에 따라 이미 관련된 많은 연구가 진행되었으며 현재 웹과 관련된 음성인식시스템에 대한 연구가 활발하게 진행 중에 있다. 이러한 연구들 중에서 1998년에 연구된 '웹 브라우저 상에서 한국어 음성인식을 이용한 정보 검색 시스템'은 EVANS(ETRI Voice Activated Navigation System)을 개발하여 인터넷을 통해 음성명령을 제공하였다. 2001년에 발표된 '한국어 음성 웹 브라우저 설계 및 구현'에서는 별도의 음성 데이터베이스를 구축하여 음성 데이터를 등록하였다. 웹사이트 의 로그인 과정이 음성으로 처리되는 '웹 기반의 화자확인시스템 설계에 관한 연구'에서는 인터넷을 통해 별도의 화자인식시스템에서 음성을 학습시키고 음성 데이터를 별도의 음성 데이터베이스에 저장한다. 이와 같이 기존의 연구들은 많은 비용과 시간을 투자하여 각각의 웹사이트에 음성 데이터베이스를 구축하고 인식과정을 수행해야 하는 단점을 지니고 있다. 이러한 단점을 보완할 수 있는 방법으로 본 연구에서는 웹서비스 기술을 이용한 음성인식시스템을 구현하고자 한다.

웹서비스 기술은 표준프로토콜인 SOAP을 이용하여 응용프로그램의 요청에 따라 원격 시스템에서 특정 기능을 수행하고 그 결과를 응답해 주는 서비스이다. 웹서비스를 요청하고 그 결과를 수신 받는 응용프로그램을 웹서비스 클라이언트라고 하며 이 웹서비스 클라이언트에는 콘솔이나 윈도우 기반의 응용 프로그램과 비즈니스 로직을 포함한 웹사이트 서버가 포함된다. 이와 같이 다양한 웹서비스 클라이언트를 지원하는 웹서비스 기술은 웹서비스 클라이언트의 플랫폼과 운영체제에 독립적이므로 각각의 웹서비스 클라이언트가 많은 비용과 시간을 투자하여 별도의 인식시스템이나 음성 데이터베이스를 구축해야 하는 단점을 보완할 수 있다.

본 연구에서 구현된 음성인식시스템은 코드북과 VQ인식과정을 이용하여 인식을 수행하는 화자 독립적인 시스템이다. 코드북은 음성을 인식할 때 필요한 음성의 정보로써 입력된 음성과 비교하는 표준 음성 패턴이다. VQ인식과정에서는 특징분석프로세서를 이용하여 음성신호에서 특징벡터를 추출 후 인식할 음성정보를 포함한 코드북과 비교하여 가장 오차가 적은 코드북의 인덱스 값을 출력한다. 구현된 음성인식시스템은 차후에 한글 단어로 인식 대상을 확장하기 위하여 한글 숫자음을 인식대상으로 한다.

본 연구에서 구현한 시스템은 음성명령이 가능한 웹브라우저와 이 웹브라우저를 동작시키기 위한 웹서버, 마지막으로 웹서비스로 만들어진 음성인식시스템으로 구성되어 있다. 이 시스템의 기능은 다음과 같다.

첫 번째로 음성명령이 가능한 웹브라우저는 음성녹음 기능과 데이터 전송을 위한 ActiveX 컨트롤을 포함하고 있다. 두 번째 구성 요소인 웹서버는 웹브라우저로부터 전송되는 음성데이터를 수신하여 음성인식시스템을 제공하는 웹서비스에 데이터를 전송한다. 음성인식시스템으로부터 반환 받은 음성인식결과를 해당되는 html 페이지로 전환하여 웹브라우저로 전송한다. 마지막으로 웹서비스를 위한 음성인식시스템은 음성 데이터를 분석하고 코드북과 VQ인식기를 이용하여 음성을 인식한 후 인식결과를 웹서버에게 전달한다.

본 연구에서는 구현된 시스템의 성능을 알아보기 위하여 '영'부터 '구'까지 숫자음에 대한 링크가 포함된 html 페이지를 작성하였다. 또한 각 숫자음에 해당하는 음성신호 3-6개로 코드북을 생성하고 각 음에 대한 음성신호를 입력받아 숫자음을 인식하는 프로그램을 구현하였다.