

웹환경에 기반한 음성인식시스템 구현

오지영, 오신영, 김윤중

oz1115@empal.com, syndrom00@hanmail.net, yjkim@hanbat.ac.kr

국립 한밭대학교

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

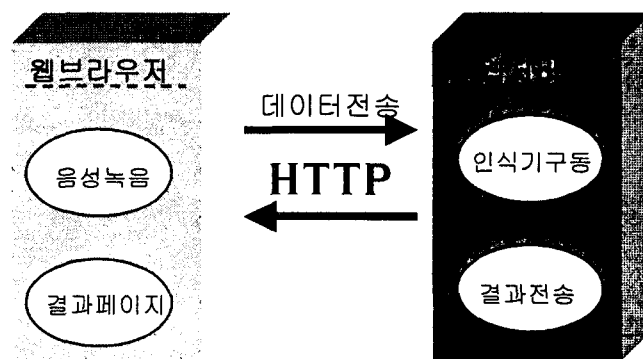
Keyword : 음성인식, 웹서비스, 웹서버

월드와이드웹(WWW)이 1989년에 등장한 이래, 비약적인 발전을 거듭하여 음성인식 기술을 적용하고자 하는 욕구가 증대되고, 음성인식기술이 이용되는 단계에 이르고 있다. 음성인식기술이 웹 환경에 적용되기 위해서는 음성인식 웹브라우저가 필요하다. 음성인식 웹브라우저는 명령을 마우스로 클릭하는 대신 음성으로 발음하여 처리될 수 있어야 한다.

음성인식 웹브라우저는 음성을 어디에서 분석하는가에 따라 다음과 같이 2가지로 분류 할 수 있다. 첫 번째, 음성이 입력되면 웹 브라우저가 음성데이터를 분석하고 분석결과가 서버로 전송된다. 두 번째, 입력된 음성데이터가 HTTP 프로토콜에 의하여 웹서버로 전송되면 웹서버가 음성데이터를 분석한다. 첫 번째의 웹브라우저는 전송량이 적어 전송속도가 빠른 장점이 있으나 브라우저에 부하가 많이 걸리는 단점이 있고 두 번째 웹브라우저는 적은 부하와 첫 번째 웹브라우저보다 다소 느린 전송속도를 갖는다.

본 연구에서는 부하가 적은 두 번째 웹브라우저를 채택하여 다음과 같은 기능을 갖는 웹환경에 기반한 음성인식시스템을 구현하고자 한다. 1)하이퍼링크를 마우스로 클릭하는 대신 음성으로 발음하여 처리 될 수 있으며 2)웹서버에서 별도의 인식기를 구동하여 인식 처리 된 결과에 해당하는 HTML 페이지를 반환한다.

본 연구에서 구현된 웹환경에 기반한 음성인식시스템은 웹 브라우저 기능, 데이터 전송기능, 웹서버기능으로 구성되었으며 상세 기능은 다음과 같다.



<그림1 시스템의 전체구성도>

1) 웹브라우저 기능 : 웹브라우저에서 하이퍼링크의 텍스트를 음성으로 녹음하여 웹서버로 전송하고 해당되는 HTML 페이지를 수신하여 웹 브라우저에서 출력한다.

① 음성녹음 : 음성녹음 기능은 Windows의 멀티미디어 자원을 지원하는 Visual C++ 언어를 이용하여 ActiveX 컨트롤로 구현하였다.

- ② 결과페이지 : 웹 서버에서 전송된 HTML 페이지를 로딩시킨다.
- 2) 데이터 전송 기능 : 녹음된 음성데이터는 HTTP 프로토콜로 웹서버에 전송된다. 웹브라우저에서 웹 서버로의 음성 데이터 전송은 ASP.NET의 파일 업로드 기능을 이용하였다. 웹 서버에서 처리된 결과(HTML 페이지)도 HTTP 프로토콜에 의하여 웹브라우저로 전송된다.
- 3) 웹서버 기능 : 웹서버 기능은 웹브라우저로부터 전송되는 음성데이터를 수신하여 인식된 결과의 해당 HTML 페이지를 찾아서 웹브라우저로 전송한다.
- ① 인식기 구동 : 수신된 음성데이터를 음성인식기로 전달하고 인식된 결과를 받아내는 시스템이다. 이 시스템은 닷넷의 웹서비스 기술로 구현하였다.
- ② 결과 전송 : 음성 분석 결과인 HTML 페이지를 웹 브라우저로 전송한다
- 본 연구에서는 웹환경에 기반한 음성인식시스템의 중요 기능인 하이퍼링크 대신 음성으로 페이지를 로딩하는 시스템을 구현하였다.
- 구현된 웹환경에 기반한 음성인식시스템의 성능을 알아보기 위하여 0부터 9까지 링크가 포함된 HTML 페이지를 작성하였다. 하이퍼링크 대신 음성데이터로 실험한 결과 음성데이터가 웹서버에서 분석되고 결과에 해당하는 HTML 페이지가 웹브라우저에 로딩되는 것을 확인 할 수 있었다.
- 향후 이 연구가 계속 되어 웹환경에서 음성명령이 가능한 웹브라우저와 모바일 환경에서 음성명령이 가능한 시스템을 개발할 예정이다.