

단어독립 음성인식 시스템을 위한 운용시스템 개발

°전예임*, 박상규, 구명완
한국통신 소프트웨어연구소, *한국통신 전산사업단

Development of the Operating and Management System for a Vocabulary Independent Speech Recognition System

Yame Jeun*, Sahng-Gyu Park, Myoung-Wan Koo
Software Research Lab., Korea Telecom Research Lab.
*Computer Operation & Development Center, Korea Telecom

요 약

이 논문은 현재 주식시장에 상장되어 있는 약 700개 회사의 현재주가를 음성인식을 이용하여 검색할 수 있는 대어휘, 화자 독립, 단어독립 음성인식 시스템(KT-STOCK)의 운용자를 위한 운용관리 시스템에 대해 기술하였다. KT-STOCK은 시스템의 음성안내에 따라 사용자가 전화기에 상장회사 이름을 말하면, 이 시스템은 그 회사의 현재 증권정보를 말해 준다. 이 시스템의 운용관리 시스템은 주식시장에 상장된 종목의 변화에 따라서 인식대상 단어를 추가하거나 삭제, 조회할 때 그 처리를 용이하게 할 수 있도록 구현되었다.

1. 서 론

음성인식기술을 이용한 주요한 응용 중의 하나는 전화망을 통하여 음성을 인식하여 각종 부가서비스를 제공하는 것이다. ARS(Audio Response System)은 명령어 입력을 위해 push-button 을 이용하므로 말에 비해 사용이 불편하고 부자연스러운 단점이 있다. 이런 단점을 극복하기 위해 음성인식을 이용한 대화형 음성기술(interactive voice technology)이 대두되었다. 응용으로는 미국 Ameritech과 Nynex 전화회사에서 설치한 Voice Service Node[3], 일본NTT 에서 금융업정보를 제공하는 ANSER[4] 등이 있다. 또한 캐나다의 BNR(Bell Northern Research)은 1992년 중반부터 주식시세 서비스를 시작했는데, 이용자는 이 시스템에 전화를 걸어서 종목명을 말하면 현재의 주가를 알 수 있다[2].

한국통신에서 개발한 음성인식을 이용한 증권정보 안내시스템(KT-STOCK)은 음성인식에 의한 정보검색 모델시스템 개발의 일환으로 1993년 2 월에 착수되어 1994년 6월에 개발이 완료되었다[5]. 개발된 시스템의 특징은 컴퓨터가 이용자와 말을 인식하며(인식률 92%), HMM (Hidden Markov Model)기술과 4개의 고속 디지털 신호처리 프로세서를 이용한 병렬처리방법

등의 최첨단 음성인식 기술로 구현되었다. 전화를 통하여 주식시장에 상장되어 있는 700 여 회사의 현재주가와 종합주가지수를 안내받을 수 있으며, 음성인식기술을 이용함으로써 기존 ARS 서비스의 가장 큰 문제점인 복잡한 종목코드를 암기해야 하는 불편을 해소했다. 그리고, 새로운 종목이 추가될 때도 별도의 음성인식기의 훈련과정을 거치지 않고 새로운 단어의 음소리스트만 추가시키면 바로 그 단어를 인식할 수 있는 단어독립 음성인식 시스템이다. 그런데, 개발된 시스템은 운용자측면이 고려되어있지 않아서 새로운 회사가 상장되거나 종목이 없어지거나 할 때 그 처리가 용이하지않고 일일이 수작업에 의해 처리되어야 한다. 이 과정은 전문가에 의해서만 가능한 것이므로 비전문가인 운용자들도 운용할 수 있도록 하기위해서 이 운용프로그램이 개발되었다. 워크스테이션에서 구현된 운용프로그램의 시작화면에서 운용자는 원하는 메뉴를 선택하면 그에 관련된 여러가지 과정이 자동적으로 처리된다. 2장은 개발된 단어독립 음성인식시스템의 특징에 대해 설명하였으며, 3장은 개발된 운용관리프로그램에 대해 기술하였으며, 4장에서 결론을 맺었다.

2. 단어독립 음성인식시스템의 특징

기존의 음성인식시스템은 인식대상단어가 고정되어 있기때문에, 여기에 새로운 인식대상단어를 추가시키려면 그림 1과 같이 음성인식기를 다시 훈련시켜야 한다. 이를 위해서 훈련용 음성데이터를 수십명에서 수백명분을 수집해야 하는 번거로운 과정을 거쳐야 한다. 그러나, 개발된 음성인식시스템은 그림 2와 같이 기존에 가지고 있는 음성인식데이터를 가지고 새로 추가되는 단어에 대한 음소리스트를 생성시켜서 훈련과정을 거치지 않고 음성인식 대상단어에 바로 추가시킬 수 있다.

그림 3은 음성인식대상단어 리스트의 모양과 '홍길동사'라는 종목이 추가될 때를 보여주고 있다. 새로 추가될 단어의 음소리스트 'ㅎ + ㅏ + ㅓ + ㅇ + ㄱ + ㅣ + ㄹ + ㅓ + ㅓ + ㅣ + ㅓ + ㅓ'를 인식대상단어 리스

단어독립 음성인식 시스템을 위한 운용시스템 개발

트에 추가시키지만 하던 별도의 훈련과정이 필요없이 이 단어를 인식할 수 있게 된다. 그러나, 개발된 음성인식시스템은 음운현상을 고려하여 문맥중속 음소모델링을 하므로 입력문자로부터 문맥중속 음소로 변환시켜주는 프로그램이 필요하다.

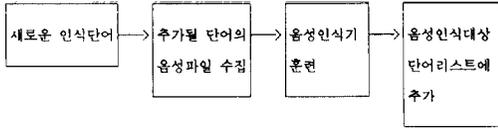


그림 1. 기존 시스템의 인식대상단어 추가과정

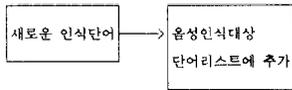
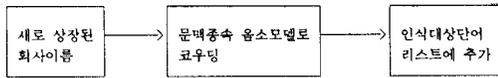
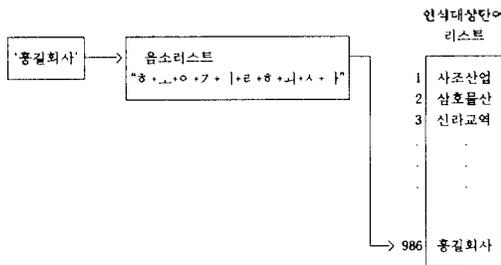


그림 2. 개발된 시스템의 인식대상단어 추가과정



a) 인식대상단어 추가



b) '홍길회사'를 추가하는 실제 예

그림 3. 인식대상단어의 추가

3. 운용관리시스템 구성

운용관리프로그램은 그림 4와 같이 현재가 DB처리 모듈, 인식대상단어 처리 모듈, 음성파일 녹음 모듈, PC interface 모듈 등으로 구성되어 있다. 현재가 DB는 증권정보안내시스템에서 서비스해 주는 실제 증권정보들이 들어있는 부분으로 회사이름, 현재 가격, 전일대비 가격, 동락부호, 전일대비 가격변동액,

매도호가, 매수호가, 현재 거래량 등으로 구성되어 있다. 현재가 DB처리 모듈은 새로운 종목이 추가되거나 삭제될 때 위의 정보들을 추가하거나 삭제한다. 운용관리 프로그램의 메뉴에서 '종목추가'라는 메뉴를 선택하면 새로 추가될 회사의 code가 나타나고 회사이름을 입력하면 현재가 DB에 새로운 자리를 만들고 관련 정보들을 초기화 시킨다. '종목삭제'를 선택하면 삭제할 회사이름을 입력받고 그 회사의 현재가 DB를 없애게 된다.

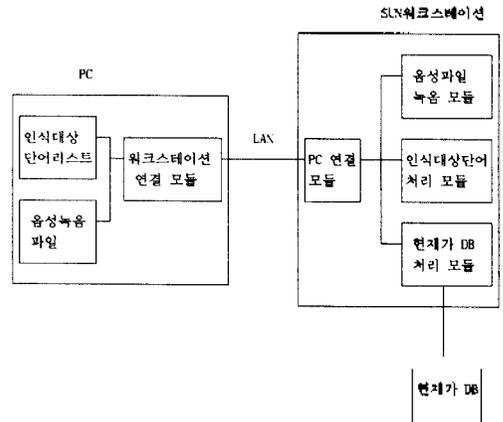
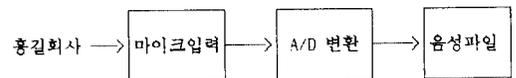


그림 4. 운용관리프로그램의 구성도

인식대상단어 처리 모듈은 음성인식부에서 사용하는 인식대상 리스트를 다루는 부분으로 추가된 종목에 대한 인식 index와 음소구성표를 만들어서 그림 3과 같이 리스트에 추가하고, 삭제된 종목에 대한 인식 index와 음소구성표를 제거한다.

음성파일 녹음 모듈은 증권정보안내시스템의 음성합성부에서 사용될 실제 목소리를 녹음하는 부분으로 새로 추가되는 종목의 이름을 아나운서가 말하면 그 소리를 A/D변환하여 파일로 저장한다. 그 과정은 그림 5와 같다.



파일이름 : C986

그림 5. 음성파일 녹음

증권정보안내시스템은 H/W 측면에서 보면 워크스테이션 부분과 PC부분으로 구성되어 있는데, 이 운용프로그램은 워크스테이션에서 수행되고 음성인식부와 음성합성부는 PC부분에 있으므로 위에서 생성한 데이터들은 PC로 전송되어야 한다. PC interface 모듈에서 전송을 담당하고 있는데, PC NFS를 이용하

여 구현되었다. 전송해야 할 데이터는 인식대상 단어 처리 모듈에서 생성한 인식대상 단어 리스트와 음성파일 녹음 모듈에서 받은 음성녹음 파일이다.

운영관리시스템에서는 상기 음성인식을 이용한 증권정보안내시스템의 상용화를 위해, 운용자가 쉽게 증권정보안내시스템을 관리할 수 있도록 워크스테이션에 운용관리를 위한 통합패키지를 구현하였다. 새로운 종목명의 텍스트 정보 추가와 음성녹음을 쉽게 할 수 있게하고, 기존 정보의 삭제, 변경에도 편의를 기하였다.

그림 6은 운용관리시스템의 상태전이 다이어그램이다. 이 시스템은 크게 세부분으로 구성되어 있는데, 현재 상장되어 있는 회사의 이름을 알아볼 수 있는 조회부분과 새로 상장된 회사에 관한 여러가지 정보를 현재의 데이터베이스에 추가시킬 수 있는 추가부분과 주식시장에서 없어진 종목에 대한 데이터베이스를 제거할 수 있는 삭제부분으로 구성되어 있다.

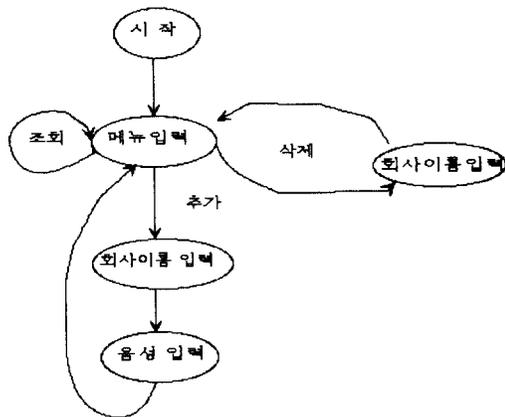


그림 6. 상태전이 다이어그램

3.1 종목조회

운영관리 시스템을 실행시키면 현재 상장된 종목명이 윈도우 상에 나열되고 스크롤바를 이용하여 원하는 종목명을 조회해 볼 수 있다. 이 때 조회되는 내용은 상장된 회사의 종목코드와 종목명이다.

3.2 종목추가

주식시장에 새로운 회사가 상장되면 그 회사이름과 종목코드를 데이터베이스에 추가시키고 인식대상 단어리스트에도 그 이름이 추가되며, 사용자에게 들려줄 음성합성용 음성파일을 추가된 종목이름에 대하여 새로 녹음한다. 종목추가 버튼을 누르면 종목코드와 종목명에 대한 텍스트 정보를 입력하게 되며, 그 다음에 종목명에 대한 음성파일을 녹음하는 프로그램이 수행된

다. 또한 새로운 종목명에 대한 음소리스트가 생성되어 PC의 음성인식부에서 사용할 수 있도록 PC로 TCP/IP socket을 통해 전송된다. 이 경우 PC의 프로그램은 클라이언트이고 워크스테이션에 있는 프로그램은 서버가 된다.

3.3 종목삭제

삭제할 종목명을 선택하면 확인 메시지가 나오고, 관련된 텍스트 정보와 음성파일, DSP 용 파일 각각에서 해당 정보는 삭제된다.

위의 처리절차 중 음소리스트 생성, 증권정보 DB update, 음성녹음 등은 이미 모듈화되어 작성되어 있으므로 여기서는 단지 이들을 결합시켜 운용자가 쉽게 관리할 수 있게 하려는 것이다. 운용자를 위한 인터페이스는 SUN 워크스테이션 상에서 Motif를 이용하여 구현하였으며, Motif는 PC 윈도우즈와 Look and Feel 이 같아서 운용자가 별 부담없이 사용할 수 있을 것이다. 개발의 보조 도구로는 X-Designer를 사용하였으며, TCP/IP 구현에는 PC/TCP develop kit을 사용하였다. 그림 7은 구현된 운용관리프로그램의 시작화면이다.

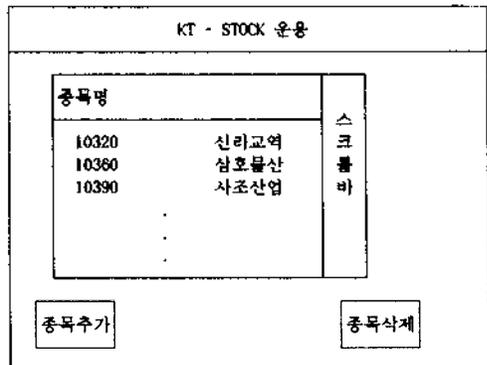


그림 7. 운용관리프로그램의 시작화면

4. 결 론

이 논문은 음성인식기술을 이용한 증권정보안내시스템의 상용화를 위해서 필요한 운용자 인터페이스를 구현하였다. 증권정보안내시스템은 이미 구축되었지만, 상용화하기에는 운영면에서 복잡하고 수작업으로 처리해야 할 일들이 많았다. 그래서, 운용에 필요한 여러 일들을 통합하여 일반 운용자도 증권정보안내시스템을 운용하기 쉽도록 하기 위해서 이 시스템을 구현하였다. 이 시스템은 현재가 DB 처리 모듈, 인식대상 단어 처리

단어독립 음성인식 시스템을 위한 운용시스템 개발

모듈, 음성파일 녹음 모듈, PC Interface 모듈 등으로 구성되어 있는데 새로운 종류가 추가될 때 그 종류에 대한 훈련은 시킬 필요가 없고 인식대상단어의 음소리스트만 추가하면 그 종류를 인식할 수 있게 되었다. SUN 워크스테이션에서 Motif를 이용하여 구현하였으며, 운용자는 화면상의 메뉴란 눌러주면 모든 과정이 자동적으로 처리된다.

참 고 문 헌

- [1] "자동통역전화 요소기술연구", 한국통신소프트웨어연구소 '94년도 연구과제보고서, 1994.12
- [2] "자동통역전화 요소기술연구", 한국통신소프트웨어연구소 '93년도 연구과제보고서, 1993.12
- [3] M. Lenning, "Putting speech recognition to work in the telephone network," IEEE computer, vol.23,no.8, pp.35-41, Aug.1990
- [4] R. Nakatsy, "Answer: an application of speech technology to the Japanese banking industry," IEEE computer, vol.23, no.8, pp.43-48, Aug.1990
- [5] M. -W. Koo et al., "KT-STOCK: A speaker-independent, large-vocabulary speech recognition system over the telephone" in Proc. 1994 IEEE Int. Conf. Acoust., Speech, Signal Processing, Sep., 1994.