

# 기업체 음성 다이얼링 시스템으로 인식단어 다운로드 자동화 구현

°류상선\*, 김철\*, 김재인\*

\* 한국통신 멀티미디어연구소 음성언어연구실

## Automation of Downloading Recognition Words to The Business-Name Voice Dialing Service System

°Chang-Sun Ryu\*, Chul Kim\*, Jae-In Kim\*

\* Spoken Language Research Division, Multimedia Tech. Research Lab., Korea Telecom.  
csryu@kt.co.kr, kchul@smm.kotel.co.kr, jikim@smm.kotel.co.kr

### 요 약

본 논문은 화자독립 음성인식 시스템인 기업체 음성 다이얼링 시스템의 서비스 내용과 시스템 명상에 대하여 설명하고 본 시스템으로 인식단어의 추가 및 삭제의 자동화를 위해 설계 및 개발한 인식단어 적용을 위한 실행파일 생성기에 대하여 기술한다. 기업체 음성 다이얼링 시스템은 사용자가 기업체 명칭, 상품명등의 기업체를 대표하는 명칭을 말하면 해당되는 기업체로 전화를 걸어주는 서비스를 제공한다. 본 시스템의 원활한 운용을 위해서는 인식단어의 추가 및 삭제가 즉시 이루어져야 한다. 이러한 요구사항은 음성인식 시스템 운용에 있어서 대단히 중요한 과제가 된다. 그러므로 본 논문에서 기술하는 실행파일 생성기가 그러한 문제를 해결하는데 하나의 해결책이라 생각된다.

### 1. 서 론

현재 국내 음성인식 기술은 일반 이용자들에게 가깝게 느껴지고 있다. 그중 대중화에 큰 역할을 담당할 것은 이동통신 서비스 분야에서 단말기를 통한 음성다이얼링이라하겠다. 이에 반해 본 한국통신의 기업체 음성 다이얼링 서비스는 공중전화망을 통한 음성인식 서비스로서 전자는 단순한 음성인식만을 제공하는데 반해 기업체 음성다이얼링은 교환기능이 들어가는 공중전화망상에서의 서비스라 하겠다. 기업체 음성다이얼링 서비스는 음성인식을 하여 해당 전화번호로 호 전환을 하는 기능적인 측면에서 본 연구소에서 개발한 음성 다이얼링 서비스와는 같지만 가입자 측면과 인식단어 운용 측면에서 서로 다르다. 즉, 음성다이얼링 서비스의 가입자는 일반이용자로서 서비스 등록 후 인식명칭에 해당하는 전화번호를 등록하여 사용하지만 기업체 음성 다이얼링 서비스에서의 가입자는 일반 기업체이다. 기업체에서 요구하는 인식명칭과 해당 전화번호를 본 시스템에 등록하여 일반 이용자가 사용한다. 이러한 기업체 음성다이얼링 서비스를 제공하기 위해 본 시스템의 기본 구조는 크게 교환기능과 음성인식 기능을 수행하는 교환접속 장치(VDE, Voice Dialing Exchange)와 가입자 정보 등록, 음성인식 프로그램 생성 및 합성을 관리

등 가입자정보 관리 기능을 담당하는 가입자관리장치(SMC : Subscriber Management Controller)의 2 부분으로 나누어지었다. 기존의 음성다이얼링 시스템은 인식단어 수가 150 단어이며 인식단어는 고정되어 바뀌지 않으므로 인식단어의 추가 및 삭제 작업이 필요치 않았다. 그러나 기업체 음성다이얼링 서비스는 인식단어의 추가 및 삭제가 가입자의 요청에 의해 발생하므로 실시간으로 인식단어를 갱신하여 교환접속장치로 다운로드 하여야 한다. 그러므로 기존 서비스의 원활한 운용을 위해 인식단어의 추가 및 삭제등의 작업이 필요하다. 그러므로 본 논문에서는 기업체 음성 다이얼링 시스템의 서비스 내용과 시스템 명상에 대하여 설명하고 본 시스템으로 인식단어의 추가 및 삭제의 자동화를 위해 설계 및 개발한 인식단어 적용을 위한 인식프로그램을 생성하는 실행파일 생성기(이하 : 인식단어 실행파일 생성기)에 대하여 기술한다.

### 2. 서비스 개요 및 시스템 형상

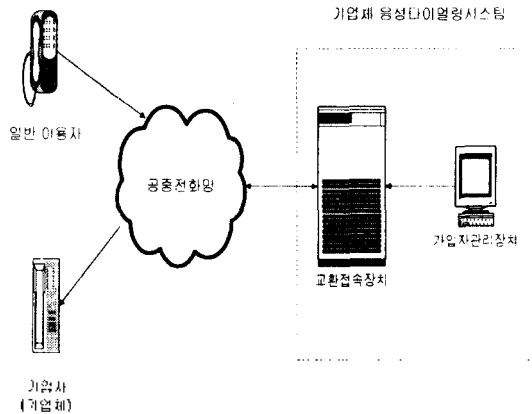
본 장에서는 기업체 음성다이얼링시스템의 기본구조 및 서비스 시나리오와 인식단어의 추가 및 삭제등의 작업의 필요성 및 기존의 방법에 대하여 살펴본다.

#### 2.1 시스템의 기본 구조 및 시나리오

##### 가. 시스템 기본 구조

기업체 음성다이얼링 시스템(BVDS: Business-name Voice Dialing System)은 공중전화망에서 사용되고 있는 국선 교환기(TDX-10)를 기본으로 개발된 일반 가입자 음성 다이얼링 시스템의 H/W를 근간으로 하여 교환접속장치를 구현하였으며 서비스 변경에 따른 호 처리 및 운용, 유지보수를 위한 S/W를 새로이 설계 및 개발되었다. 이미 앞에서 설명하였듯이 기업체 음성 다이얼링 시스템은 교환기능 및 음성인식 기능을 수행하는 교환 접속 장치와 가입자 정보등록, 음성인식 프로그램 생성 및 합성유이 관리등을 하는 가입자관리장치로 구분된다 [1].

자는 PC에서 인식프로그램을 실행하여 ftp를 이용 가입자 처리장치로 전송을 한다. 가입자 관리장치에서는 인식프로그램을 수신하여 이를 다시 교환접속장치의 VRU로 ftp를 통해 다운로드하는 일련의 과정을 수작업으로 수행하여야 한다.

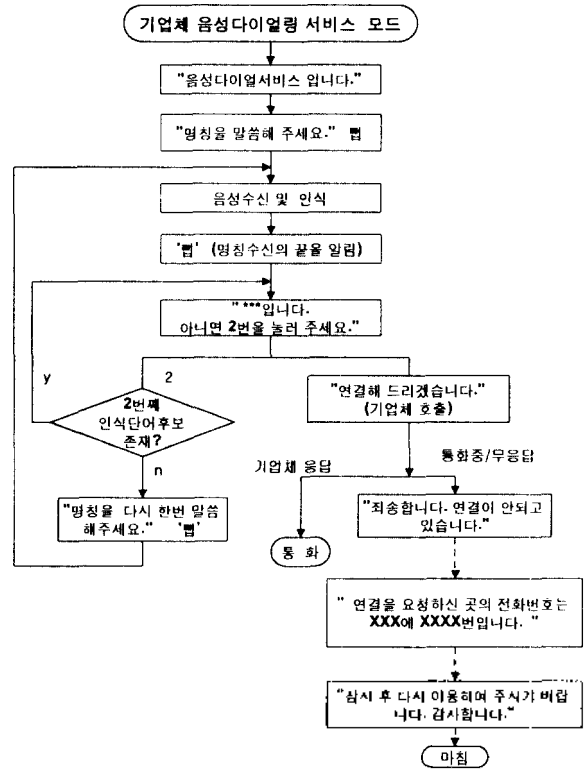


<그림 1> 기업체 음성다이얼링 시스템 개략도

<그림 1>은 본 기업체 음성다이얼링 시스템의 개략도이다.

나. 서비스 시나리오

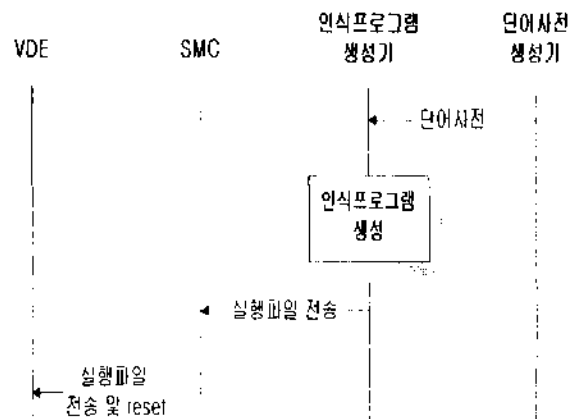
기업체 음성다이얼링서비스를 제공하기 위한 기본 서비스 시나리오는 다음과 같다. 가입자가 전화망을 통하여 서비스 시스템의 접속번호를 누르면 시스템에 연결되고 시스템은 통화하고자 하는 기업체의 명칭을 음성으로 입력할 것을 요구한다. 가입자가 안내방송에 따라 인식명칭을 입력하면 시스템은 음성인식 후 인식결과에 대해 가입자에게 확인을 요구하고 맞으면 기업체 데이터베이스에서 착신전화번호를 추출하여 자동으로 다이얼링하여 연결을 시켜준다. 이때 착신 기업체가 통화중 또는 부재중이어서 통화가 완료되지 못하면 해당 전화번호를 음성으로 안내하여 준다<그림2>[1].



<그림 2> 서비스 시나리오 개략도

2.2 인식 프로그램 생성기

음성인식 서비스 제공을 위해 인식 프로그램은 교환접속장치(VDE)의 VRU(Voice Recognition Unit) 모듈의 DSP 메모리에 상주하게 된다. 인식프로그램은 인식을 위해 필요한 데이터와 인식 알고리즘으로 크게 두 부분으로 분류할 수 있다. 이미 앞에서 언급하였듯이 기존 음성다이얼링 서비스 제공에 있어서는 인식프로그램의 업그레이드 작업 이외에는 VRU의 인식프로그램을 관리할 필요가 없다. 그러나 기업체 음성 다이얼링 서비스는 인식단어의 수정이 계속되므로 이에 대한 자동화가 필요하다. 그러므로 본 절에서는 기존의 인식프로그램 생성 후 시스템 적용 방법에 대해 살펴본다. <그림 3>에서 보는바와 같이 인식프로그램은 외부 장치인 DSP 모듈이 장착된 PC에서 만들어져서 VDE로 다운로드 된다. 기존에는 인식프로그램의 수정이 없으므로 시스템 운용



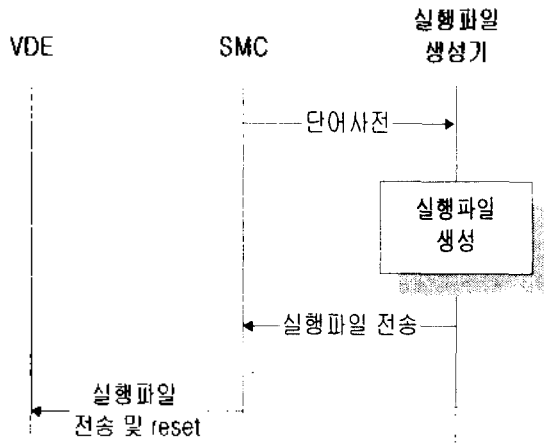
<그림 3> 기존 인식프로그램 생성 및 적용방법

## 기업체 음성 다이얼링 시스템으로 인식단어 다운로드 자동화 구현

### 3. 인식단어 실행파일 생성기

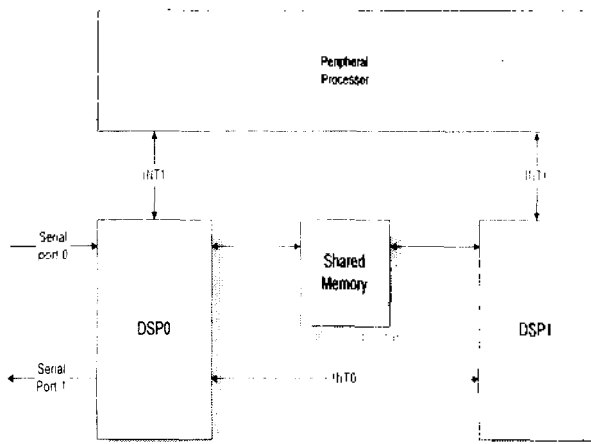
#### 3.1 개요

인식프로그램 생성의 자동화를 위해서는 시스템 운영자로 하여금 외부 혹은 내부에 있는 장치를 이용하길간에 가입자관리장치(SMC)에서 임무를 처리할 수 있어야 한다. 그러기 위해서는 앞장에서 언급한 인식프로그램 생성기와 가입자관리장치간 통신이 이루어져야 하며 <그림 4>는 가입자관리장치와 인식단어 실행파일 생성기 간 업무처리결차를 보여준다.



<그림 4> 인식단어 실행파일생성기 작업흐름

인식프로그램은 두부분으로 나누어져 있다. 그래서 음성인식을 위해 두개의 DSP를 사용한다. 특징추출과 벡터 양자화(Vector Quantization)를 담당하는 DSP0와 비터비 검색(Viterbi Search)를 하는 DSP1이 있다. <그림 5>는 기업체 음성다이얼링 시스템의 음성인식부 H/W 구성을 보여준다[2].



<그림 5> 음성인식부 H/W 구성

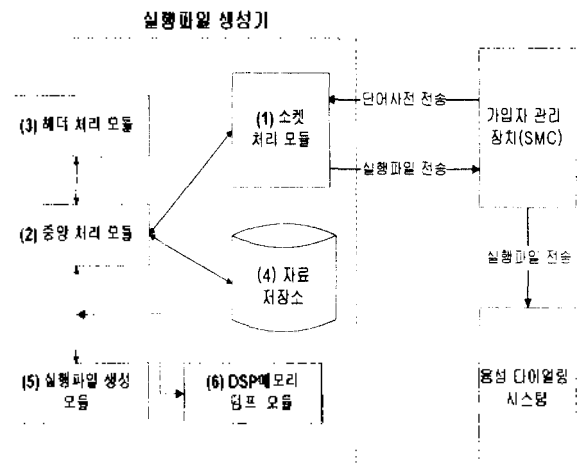
#### 3.2 시스템 구현

##### 가. 시스템 요구사항

자동화된 인식단어 실행파일 생성기의 구현을 위한 요구사항은 다음과 같다. <그림 3>에서의 단방향 작업흐름에서 <그림 4>으로의 양방향 작업형태로 하기위해 UNIX 기반의 가입자관리장치와 DOS 기반의 실행파일 생성기간의 이기종간 통신을 해야한다. 그리고 인식프로그램은 2개의 DSP를 사용하지만 인식단어 변경에 해당되는 부분은 DSP1만 해당되므로 인식단어 실행파일 생성기능을 수행하도록 1개의 DSP 보드로도 가능하다.

##### 나. 시스템 구현

가입자관리장치와 인식단어 실행파일생성기간 통신을 하며 인식단어 실행파일생성 작업을 수행하기 위해서는 실행파일생성기의 개발환경을 우선 고려하여야 한다. 첫째, 인식알고리즘을 위한 DSP 컴파일러가 있고 둘째, 초기화 작업을 위한 Borland C++ 컴파일러 그리고 통신을 위한 PC/TCP가 있다. 그리고 가입자관리장치와 인식단어 실행파일 생성기는 TCP/IP의 Socket을 사용하는 클라이언트-서버 모델을 따른다. socket을 이용하여 통신 프로그램 작성시 socket을 이용하는 자재는 그리 어렵지 않다. 그러나 운영체제를 달리하는 인식단어 실행파일 생성기인 PC와 통신을 하기위해 소프트웨어적으로 PC/TCP를 인스톨하여야 한다[3]. PC/TCP와 UNIX의 TCP/IP와는 흡사하지만 PC/TCP와 어떻게 접촉을 시키야 하는지가 프로그램 구현상 문제라 할 수 있다. 그러므로 PC에서 PC/TCP의 socket을 이용하는 함수들은 하나의 파일로 구성하여야 한다. 그 이유는 Borland의 include 디렉토리 파일들을 사용하는 경우와 PC/TCP의 include 디렉토리를 사용하는 경우로 구분되기 때문이다. 그러므로 잘못 include 할 경우에는 개발 스펙에 맞는 결과를 얻을 수 없다[4][5].



<그림 6> 인식단어 실행파일 생성기 구조

<그림 6>은 음성 다이얼링 서비스 시스템에서 인식단어 실행파일 생성기의 구조도 및 실행파일 생성기와 직접

적인 인터페이스를 갖는 가입자 관리장치(SMC)간 관계를 보여주고 있다. 가입자관리장치는 인식명칭이 포함된 단어사전 파일을 생성하여 인식단어 실행파일생성기로 전송을 하고 인식단어실행파일생성기로부터 실행파일을 수신하여 교환접속장치로 다운로드한다. 인식단어 실행파일 생성기는 단어사전 수신 및 실행파일 전송을 담당하는 소켓 처리 모듈(1)과 수집된 자료에 대해 수행할 작업들을 제어하는 중앙처리 모듈(2)에서는 시스템간 통신을 하므로 헤더에 대한 처리모듈(3) 그리고 수집된 단어사전 및 실행파일을 저장하는 자료 저장소(4) 자료저장소에 저장된 신규 단어사전을 이용하여 음성인식 프로그램과 인식단어가 포함된 실행파일을 생성하는 실행파일 생성 모듈(5)과 실행파일 생성모듈을 통해 DSP 보드의 메모리에 업로드된 실행 파일의 자료저장소의 파일로 다운로드하는 DSP 메모리 덤프 모듈로 구성된다. 이러한 각 모듈의 실행이 완료된 후 가입자 관리장치는 서버인 실행파일 생성기로부터 교환접속장치로 다운로드하여 사용할 수 있는 실행파일을 수신하게 된다.

### 3.3 시스템 적용 및 시험 결과

가입자관리장치에서 인식단어 처리기능이 개발중에 있으므로 SUN 워크스테이션에 인식단어처리 기능을 수행하는 시험용 클라이언트 프로그램을 개발하여 본 논문에서 기술한 인식단어 실행파일 생성기를 수행하여 교환접속장치로 전송 다운로드하여 시험하였다. 시험을 위해 사용한 데이터로 기존의 음성 다이얼링서비스는 약 130개의 문맥종속 음소를 사용하지만 기업체 음성 다이얼링 시스템에서는 약 300개의 문맥종속 음소를 사용하였으며 인식단어수는 150개로 동일하며 인식명칭은 기업체 명칭으로 변경 하였다. 시험한 결과 인식률 및 속도에 있어서는 기존의 방법에 따라 수행한 결과와 같은 결과를 얻었다. 그러나 실행파일 생성기에서의 실행파일 생성시간이 약 2배정도 증가하였다. 그 이유는 사용되는 음소의 개수가 증가함에 따라 단어사전 추가화에 소요되는 시간이 그만큼 증가하는것으로 생각된다. 이 데이터 초기화의 시간을 줄이기 위해 기존의 ASCII 로 된 음소들을 현재는 어진 데이터 포맷의 문맥종속 음소를 이용할 수 있도록 하므로써 초기화에 소요되는 시간을 단축하였다.

## 4. 결론

기업체 음성다이얼링 시스템은 인식명칭으로 기업체명 혹은 기업체를 대표할 상표명등을 인식하여 등록된 전화번호로 호전환을 하여주는 서비스를 제공한다. 그리고 서비스 운용에 있어 인식명칭의 변경이 발생하는 특징을 갖고 있다. 그러므로 본 논문은 시스템 운용을 위한 자동화된 인식단어 실행파일 생성기에 대하여 설명하였다. 시험 결과 인식률 및 인식속도는 만족스러웠으며, 초기화에 걸리는 시간은 2 배정도 증가하여 기존 300개의 ASCII 문맥종속 음소를 3개의 어진 데이터 포맷의 문맥종속 음소로 전환 사용하므로 초기화 소요 시간을 만족시켰다. 향후 변경되는 음소만 초기화하는 시스템 형상으로 전환되어야 한다고 사료된다.

## 참고문헌

- [1] 김문식,김철,김희경,“기업체 음성다이얼링 서비스 구현”,제3회 통신 소프트웨어 학술대회,141~144쪽,1998.7.
- [2] 박상규,“음성다이얼서비스 시스템의 음성인식부”,기술이전자료,한국통신 멀티미디어연구소,1997.7.10.
- [3]“PC/FCP Network Software for Dos” User’s Guide, ftp Software.Inc.1993.
- [4] 류상선,“Socket 을 이용한 PC 와 UNIX 간 통신”,TM, 한국통신 멀티미디어연구소,1998.5.
- [5] Stevens,“UNIX Network Programming”,Prentice-Hall,1990.