

호출 명령어 방식 핵심어 검출 시스템의 임베디드 DSP 구현에 관한 연구

송기창[†], 강철호^{**}

요 약

최근 핵심어 검출 시스템은 유비쿼터스 홈네트워크의 UI(User Interface)기술로써 각광받고 있다. 핵심어 검출 시스템은 TV, 라디오, 떠드는 소리 등과 같은 동적 생활 잡음에 매우 취약하다. 특히, 실제 임베디드 DSP(Digital Signal Processor) 환경에서는 상대적으로 CPU(Central Processing Unit) 연산능력이 떨어지므로, 실시간으로 입력되는 음성을 인식하기가 어려워 인식율은 급격히 하락하게 된다. 본 논문은 임베디드 DSP 환경에서 원활한 연속음성인식을 수행하기 위하여 '나래야', '홈매니저' 등과 같은 호출명령어를 선정하고 잡음을 포함한 묵음구간과 호출명령어로 구성된 최소의 인식네트워크를 토큰으로 구성하여 입력된 음성 에 대해 실시간 음성인식을 계속적으로 수행한다.

A Study on Embedded DSP Implementation of Keyword-Spotting System using Call-Command

Ki-Chang Song[†], Chul-Ho Kang^{**}

ABSTRACT

Recently, keyword spotting system is greatly in the limelight as UI(User Interface) technology of ubiquitous home network system. Keyword spotting system is vulnerable to non-stationary noises such as TV, radio, dialogue. Especially, speech recognition rate goes down drastically under the embedded DSP(Digital Signal Processor) environments because it is relatively low in the computational capability to process input speech in real-time. In this paper, we propose a new keyword spotting system using the call-command method, which is consisted of small number of recognition networks. We select the call-command such as 'narae', 'home manager' and compose the small network as a token which is consisted of silence with the noise and call commands to carry the real-time recognition continuously for input speeches.

Key words: Call-Command(호출명령어), Keyword Spotting(핵심어검출), Embedded DSP(임베디드 DSP)

1. 서 론

최근 홈네트워크 분야는 매년 30만호에 달하는 신규 아파트 및 주택을 대상으로 과거의 단순한 홈오트

메이션시장에서 각종 가정내 전동제어 및 냉난방공조시스템, 가전제품의 제어를 통합하고 초고속 인터넷을 이용하여 각종 정보를 손쉽게 접근할 수 있고 집밖에서도 집안내의 시스템을 제어할 수 있는 지능

* 교신저자(Corresponding Author) : 송기창, 주소 : 서울특별시 송파구 신천동 17번지 잠실파크리오 아파트 115동 3501호(138-240), 전화 : 02)2109-6740, FAX : 02)2109-6741, E-mail : kcsong78@hanmail.net
접수일 : 2010년 2월 2일, 수정일 : 2010년 5월 25일
완료일 : 2010년 6월 7일

[†] 정회원, 광운대학교 전자통신공학과 박사과정, 아이폰 연구소장

^{**} 정회원, 광운대학교 전자통신공학과 교수
(E-mail : chkang5136@kw.ac.kr)

* 본 연구는 2009년도 광운대학교 교내학술연구에 의하여 수행하였음.

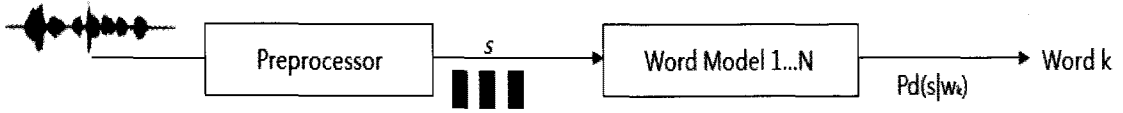


그림 1. 기존 고립단어 인식 시스템의 블록다이어그램

형 홈네트워크 시스템으로 진화시키고 있다. 따라서, 기존의 단순한 기능에서 다기능화된 시스템을 효율적이고 간편하게 접근할만한 새로운 인터페이스 기술이 요구되고 있으며 많은 건설사들과 홈네트워크 전문회사들은 이러한 인터페이스를 제공하기 위해 음성인식시스템에 대한 요구가 급증하고 있다[1-2].

기존의 홈네트워크용 월패드(wall pad)나 주방 TV폰 등에 핵심어 검출 연속 음성인식 DSP를 내장하여 거실내 공간에서의 원거리 음성인식을 수행하기 위해서는 가정환경에서 발생할 수 있는 다양한 잡음원에 대한 미등록어 거절처리를 실시간으로 수행하고, 화자의 실제 명령에 대해 인식을 수행하도록 입력신호에 대기하여야 한다. 이를 위해서 다양한 방법의 미등록어 거절 알고리즘들이 연구되어 왔다[3-6].

하지만, 상대적으로 연산처리 능력이 떨어지는 임베디드 DSP 환경에서는 화자가 발생한 핵심어 인식 결과인지, 가정환경 잡음원에서 발생한 노이즈에 의한 인식 결과인지 판별하는 연산 처리 과정이 실시간으로 이루어지기 어렵다. 오히려, 미등록어 판별 연산을 수행하는 중 화자의 실제 명령이 입력되는 경우가 많아 실제의 음성 인식이 급격히 하락할 수밖에 없다. 본 논문에서는 “나래야”, “홈메니저”와 같은 최소의 호출 명령어 인식 네트워크를 구성, 판별함으로써 300MHz CPU 성능을 갖는 임베디드 DSP 환경에서도 실시간 연속 음성인식이 가능한 시스템을 구현하였다.

본 논문의 구성은 아래와 같다. 2장에서는 기존의 핵심어 검출 시스템의 구성에 관해 설명하였고, 3장에서는 본 논문에서 제안하는 호출명령어 기반 핵심어 검출 시스템의 DSP 구현에 관해, 4장에서는 모의 실험 및 비교결과에 대하여, 5장에서는 결론 및 고찰에 대해 각각 기술하였다.

2. 기존의 핵심어 검출 시스템의 구성

음성인식 시스템은 인식하는 방식에 따라 다음과

같이 고립단어 인식 시스템과 연속음성 인식 시스템으로 분류된다[7].

2.1. 고립단어 인식 시스템

그림 1과 같은 고립단어 인식 시스템은 적은 계산량이 요구되어, 소규모 음성인식에 적합해 완구 등에 DSP로 임베딩되어 상용화 되었다. 하지만, 음성인식을 위한 별도의 녹음 버튼이 필요하며, 미등록어에 대한 처리가 고려되어 있지 않는 불편함이 있다[8-9].

2.2 기존의 핵심어 검출 시스템

고립단어 인식 시스템의 한계를 극복하고 연속적인 음성을 인식하기 위해 연구된 핵심어 검출 시스템은 그림 2와 같다. 입력되는 음성의 특징을 추출하여 핵심단어(Keyword) 및 필러(Filler)모델이 연결된 인식네트워크와 패턴비교 및 언어모델 비교를 통하여 핵심단어를 검출함으로써, 인식어인지 비인식어인지를 판별하게 된다. 이러한 핵심어 검출 시스템은 음성인식을 하기 위한 별도의 이벤트가 필요없이 연속적으로 입력되는 음성에 대해 처리가 가능해 높은 사양의 CPU를 갖는 서버급 PC등에서 연속음성인식 시스템으로써 적합해 CTI 서버를 이용한 ARS 시스템 등에 적용되어 왔다. 그러나 많은 계산량과 대규모의 데이터 저장용량이 필요하여 내장형 단말기의 임

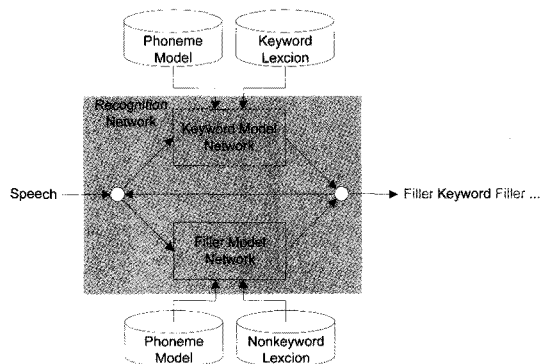


그림 2. 기존의 핵심어 검출 시스템의 인식 네트워크

베디드 환경에는 적합하지 않은 단점이 있다[10-11].

3. 제안한 호출명령어 기반 핵심어 검출 시스템의 DSP 구현

따라서, 본 논문에서는 음성 녹음을 위한 특별한 이벤트가 필요없는 기존의 핵심어 검출 시스템의 장점을 유지하면서, 임베디드 환경에서 소량의 연산으로 실시간 인식 처리가 가능하도록 다음과 같은 새로운 2단계 인식 네트워크를 제안한다.

3.1 제안한 호출 명령어 기반 인식 네트워크

기존의 핵심어 검출 인식 네트워크는 각 단어모델의 앞뒤에 묵음(silence)이 있다고 가정하고 각 단어 모델은 인식하고자 하는 하부 단위 모델의 네트워크로 구성되어 진다. 입력되는 음성 혹은 음향으로부터 추출된 특징 파라미터는 구성된 인식 네트워크의 경로를 통해 인식되어 지는데, 본 논문에서는 그림 3과 같은 2단계 인식 네트워크를 제안한다.

그림 3에서와 같이 호출 명령어에 대한 인식 네트워크는 음절별로 나누어지고 각 음절은 left-to-right 모델로 구성되어지며 필러 모델은 한 음절로 구성된

7개 이하의 음절 모델로 구성된다. 각각의 입력 음성 프레임은 단지 1개의 연속 인식 네트워크와 실시간 확률 비교가 이루어지게 된다. Token-Passing 알고리즘에 의해 디코딩하면 최대 100(10 * 10)개의 토큰만이 필요하게 되며, 인식율이 저하되지 않는 한도에서 빠른 계산을 위하여 빔폭은 40-50개의 토큰만으로 제한하여 계산량을 최소화 시킨다. 여기서 40개의 토큰이 실시간으로 계산되면서, 호출명령어 인식 네트워크에서 묵음 모델로 천이하는 순간 역전파 경로(Back Propagation Path)를 검색했을 때, 호출명령어 "나래야"로 인식되었으면 화자가 명령어를 발생하고 있다고 가정하고 연속음성인식 네트워크로 전파한다.

그림 3의 실시간 호출명령어 인식네트워크의 계산량은 실시간 MFCC 파라미터 추출 및 40 토큰 Decoding에 필요한 계산량으로 중소규모 인식기인 300단어 연속음성인식 네트워크(300*300 토큰)에 비해 극소량에 불과하다. 또한, 이론상 1000(10*10*10)번 중 1번의 오동작을 가지게 된다. 이것은 TV, 라디오, 가족간의 대화 등 Non-Stationary Noise가 항상 존재하는 가정환경에서 에너지와 ZCR(Zero Crossing Rate)를 이용하는 기존의 VAD(Voice

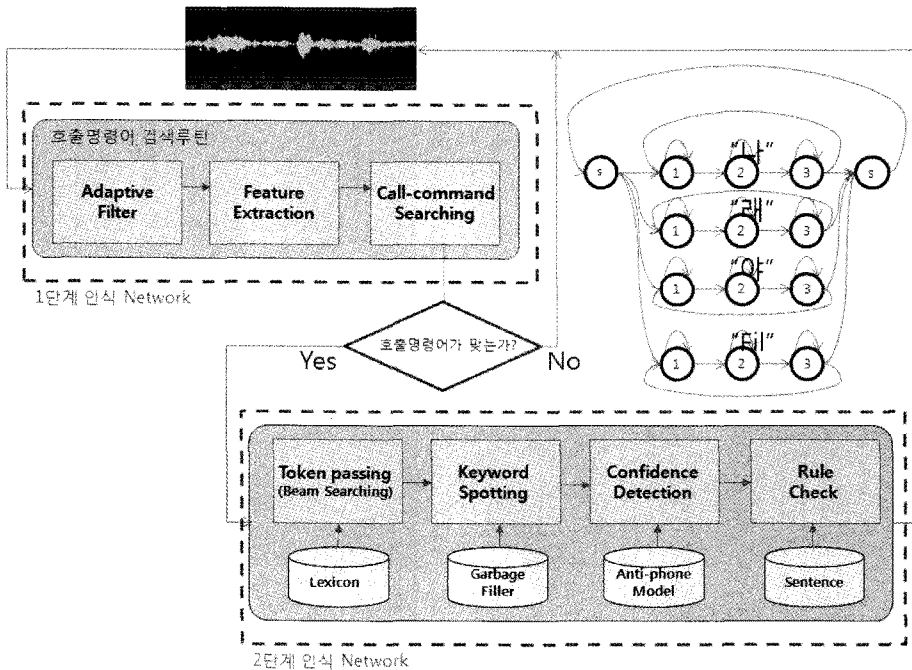


그림 3. 제안한 호출 명령어 인식 네트워크

Activity Detection)기법에 의한 오동작율에 비해 매우 우수한 성능을 보인다. 제한한 호출 명령어 인식 네트워크는 기존의 핵심어 검출 방식의 연속음성인식 기법이 방대한 계산량과 적용분야에 따라 변경하는 언어모델이나 후처리 과정을 이용해 실시간 처리가 불가능한 단점을 해결하면서, 동시에, 특별한 이벤트를 필요로 하는 고립 단어 인식시스템의 단점을 해결함으로써 음성인식을 위한 상시 대기 상태를 가능하게 하는 기법이다.

3.2 제안한 시스템의 임베디드 DSP 구현

그림 4는 복잡한 연산의 연속음성인식 알고리즘을 수행하기 위하여 Fixed Point 연산을 지원하는 TI(Texas Instrument) TMS 320VC5501을 사용한 DSP모듈의 블록다이어그램이며 그 특성은 아래와 같다[12].

- 300MHz의 내부 클럭으로 클럭당 3.33ns의 시간
- 16Kbyte의 Instruction Cache
- 16Kbyte × 16bit 의 On-Chip ROM 내장
- 4M byte의 Nor-Flash ROM
- 최대 8M × 16bit SDRAM 외부 메모리 공간
- 6개의 DMA(Direct Memory Access) 지원
- 2개의 McBSP(Multichannel Buffered Serial Port) 지원
- 4개의 타이머 지원
- I2C 지원
- 8 bit HPI(Parallel Host-Port Interface) 지원
- 단일 채널 UART(Universal Asynchronous Receiver) 지원
- 176핀의 LQFP(Low-Profile Quad Flatpack) 타입
- 1.26V의 Core Supply Voltage
- 3.3V의 I/O Supply Voltage

본 논문에서 화자독립 연속음성인식 엔진의 DSP 구현을 위하여 음성의 입력에 대한 음성신호 검출 동작과 인식동작을 독립적으로 설계하여 실시간 음성 처리를 기반으로 설계하였다. 또한, 홈네트워크의 다양한 응용시스템에 임베딩하기 위한 DSP모듈의 외부 인터페이스를 그림 5와 같이 표준화하여 설계하였다.

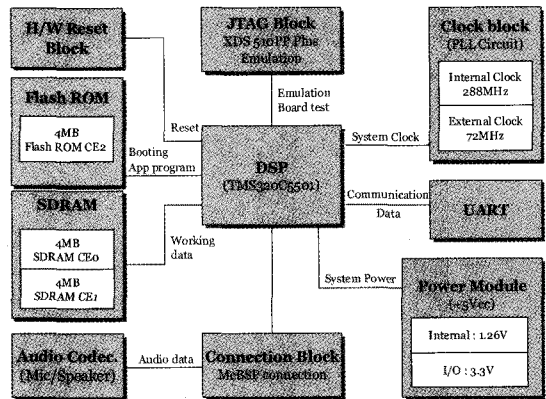


그림 4. 제안한 연속 음성인식 DSP 모듈의 블록다이어그램

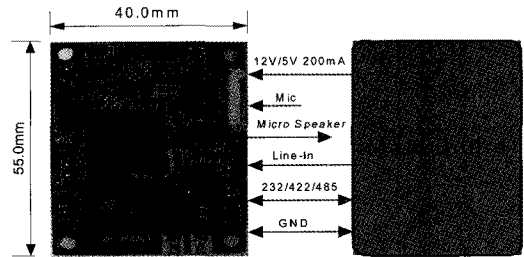


그림 5. 홈네트워크 응용을 위한 구현된 DSP모듈의 외부 인터페이스

또한 음성 Data를 입력받기 위한 회로로 16KHz의 샘플링과 16bit의 data로 변환하는 ADC와 1W의 출력을 가지는 D-Class Amp를 가지고 있는 WM8960을 Codec로 사용하였다. Codec은 DSP_Clk에 의하여 Master Mode로 동작하게 설정되었으며 입력되는 음성을 디지털화하기 위한 모든 신호를 Codec 자체에서 발생하고 DSP가 Slave로 동작하도록 System을 설계하였으며 Codec에서 처리된 음성 data는 I2S 방식의 Interface를 이용하여 DSP로 전달 되어서 처리 결과에 대한 음성안내는 DSP에서 Codec으로 전달되어 Codec내부의 DAC를 통하여 내장된 Amp에 의해 스피커를 통하여 외부에 전달되어 지도록 설계하였습니다.

S/W 개발환경으로는 TI사의 CCS 개발자 Version을 이용하였다. TI사에서 제공하는 각종 Lib에 대한 참고하였고 USB type JTAG를 이용하여 Debug 및 초기 Nor-Flash ROM을 programming을 하였으며 H/W의 회로도 설계에는 Or-Cad와 PCB Cad는 PADS를 사용하여 설계 하였다.

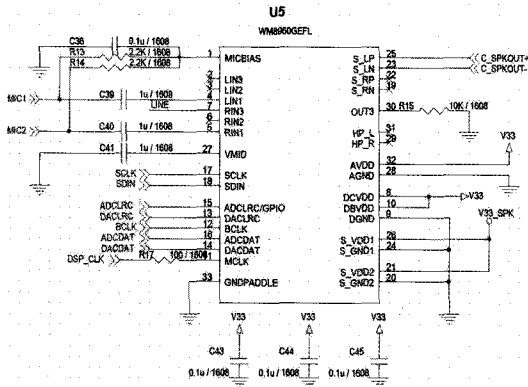


그림 6. Codec 인터페이스 회로 및 D-Class Audio 앰프

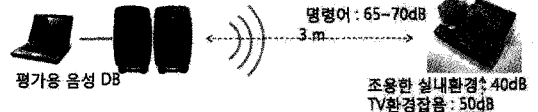
4. 실험 및 비교 결과

4.1 실험 환경

본 논문의 훈련에 참여한 음성 데이터 베이스는 화자독립형 원거리 인식을 위하여 성별, 연령별, 지역별로 홈네트워크에 사용되는 545문장을 남녀 각 1,000명씩 구축하였으며, 30msec 구간을 1 프레임으로 분석하여 Energy, 12 MFCC(Mel-scaled Frequency Cepstral Coefficient), Delta-MFCC의 26차 특징 파라미터를 추출하였고, 추출된 파라미터를 이용하여 EM(Expectation-Maximization) 알고리즘에 의해 tri-phone을 인식단위로 모델링하였다. 각 tri-phone은 3상태의 left-to-right 천이를 가지는 상태 모델로 훈련하였다. 각 상태는 계산량 대비 우수한 성능을 나타낸 2-mixture 모델로 구성되었다.

본 논문에서 제안한 방법 및 구현된 DSP 모듈의 성능평가를 위하여 남성 및 여성 화자 100명씩의 Test DB를 구축하였다. Test DB는 실제 홈네트워크에서 사용되는 명령어들로 호출 키워드를 포함하는 200 문장을 “나래야, 안방 불켜”의 형태로 “호출 키워드(Call-Keyword) + 제어 아이템 + 명령어”로 발생되어지도록 구성하여 만들었다. 실험환경은 조용한 환경에서 노트북 상에서 16kHz 샘플링 주파수로 녹음되었고, 16bit로 양자화 하여 저장한 평가용 DB를 이용하여, 그림 7과 같이 실제 가정환경과 같이 시뮬레이션 하였다. 잡음환경은 레벨미터기를 이용하여 조용한 실험실 상태(40dB 미만), 일반 가정환경(40-50dB), TV 시청환경(50-60dB)의 3가지 상태로 분류하여 3m 거리에서 각각 테스트 하였다. 또한, 본

●정량적 시험평가 - 인식을 평가



●오동작 시험평가



그림 7. 음성 인식을 및 오동작을 시뮬레이션 평가 환경

논문에서 제안한 호출 명령어 기반 음성인식 DSP 모듈의 사용 환경평가를 위하여 남성 20명, 여성 20명 화자의 실제 환경에서의 인식을 테스트 평가를 수행하였으며, 오동작을 테스트를 수행하기 위하여 24시간 TV를 켜놓은 상태에서 DSP 모듈에 JTAG를 통해 로그를 남겨 한 달에 걸쳐 평가 하였다.

4.2 실험 결과 및 검토

표 1은 남녀 각 100 명의 Test DB에 의한 시뮬레이션 평가 결과이다. 평가 결과, 조용한 환경(40dB)에서는 기존 방식의 핵심어 검출 시스템과 제안한 방식의 시스템의 인식성능 차이가 그다지 크지 않게 나타났다. 하지만, 잡음 레벨이 올라갈수록 기존 방식의 인식성능은 급격히 하락하는 반면, 제안한 방식의 인식성능은 천천히 하락하는 것을 볼 수 있다. 또한, 여성화자의 경우 그 차이가 남성 화자보다 큰 것을 확인할 수 있다. 이는 상대적으로 음량이 적은 화자의 경우, 에너지에 크게 의존하는 기존의 인식 시스템에서 비해, 제안한 호출 명령어 방식의 핵심어 검출 시스템이 보다 효과적임을 알 수 있다. 인식속도 측면에서는 시뮬레이션 결과로부터 기존의 방식이 다소 우수한 것으로 나타났으나, 이것은 기존의 방식은 음성이 입력되는 순간부터 실시간으로 인식을 수행하지만, 제안한 방식은 호출 명령어 구간에 실시간 인식을 수행하지 못하기 때문에 나타나는 물리적 시간 차이이며, 사람이 인지하지 못할 정도의 미미한 차이이다.

표 2는 남녀 각 20명이 실제 가정환경에서 구현된 DSP 모듈을 설치하여 수시로 테스트한 결과이다. 기존 방식의 경우, 정량적 테스트 보다 실제 환경에서는 더욱 성능이 떨어지는 것을 확인할 수 있었다. 이는, TV와 같은 non-stationary noise 환경에서 테스트

표 1. 시뮬레이션 환경에 의한 평가 결과

평가항목	기존의 방식		제안한 방식		평가 방법
	남	여	남	여	
1. 음성인식율	95%	93%	95%	94%	집안환경/40dB잡음환경/3m거리
	90%	89%	93%	92%	집안환경/50dB잡음환경/3m거리
	80%	75%	89%	87%	집안환경/60dB잡음환경/3m거리
2. 오인식율	3%	3%	3%	3%	특정인식명령어 발생시 다른 명령어로 인식
3. 인식속도	1.0 sec		1.2 sec		음성입력이 끝난 상태에서 인식결과가 나오기까지의 시간차를 이용해 측정

표 2. 실제 환경에 의한 평가 결과

평가항목	기존의 방식		제안한 방식		평가 방법
	남	여	남	여	
1. 음성인식율	92%	91%	95%	95%	집안환경/40dB잡음환경/3m거리
	86%	83%	92%	91%	집안환경/50dB잡음환경/3m거리
	75%	70%	88%	87%	집안환경/60dB잡음환경/3m거리
2. 오인식율	3%	2%	2%	2%	특정인식명령어 발생시 다른 명령어로 인식
3. 인식속도	1.0 sec		1.2 sec		음성입력이 끝난상태에서 인식결과가 나오기까지의 시간차를 이용해 측정
4. 거절율(오동작율)	97%(3%)		99.9%(0.1%)		평상 대화음, TV 소리등에 의한 오동작(24시간, 1,000 번의 이벤트로 계산)

트한 경우, 기존의 방식은 이미 인식을 수행하여 화자의 음성이 녹음되지 못해, 인식율이 급격히 하락하는 것을 알 수 있다. 반면, 제안한 방식의 경우, 시뮬레이션 환경(테스트 DB)에 의한 평가결과(표 1)와 거의 동일한 수준의 평가 결과를 보였다.

미등록어 거절율(오동작율)에서도 기존의 핵심어 검출 시스템이 평균 하루에 30회 정도의 오동작 횟수(오동작율 3%)를 나타낸 반면, 똑같은 후처리 과정을 포함하고 있는 제안한 호출 명령어 방식의 연속 인식시스템에서는 평균 하루 1회 정도의 오동작(오동작율 0.1%)이 발생하였다. 이는 제안한 호출 명령어 방식의 핵심어 검출 시스템이 제한된 리소스를 갖는 임베디드 DSP 환경에서 연속 음성인식 시스템의 장점을 최대화 할 수 있음을 보여주는 결과이다. 제안한 방식은 최소의 계산량과 2M Bytes 미만의 작은 메모리 공간으로서 일반 PC에 비해 상대적으로 열악한 리소스를 갖는 임베디드 DSP 환경에서도 연속 음성인식이 가능함을 보여주는 새로운 기법이다.

5. 결론 및 고찰

현재 유비쿼터스 홈네트워크를 위한 UI 기술로서의 음성인식에 관한 많은 연구가 진행 중에 있다. 본 논문에서 제안한 방식은 기존의 핵심어 검출 시스템을 임베디드화 함에 있어 문제점이 되고 있는 연산량 대비 인식 성능에 관한 해결 방안을 제안하였다. 이 방식에서 제안한 핵심어 검출 시스템은 집안 환경에서 발생할 수 있는 주방 후드, 헤어드라이기 등의 일정한 잡음 뿐 아니라, 가족간의 대화, TV, 라디오 등 일정치 않은 진폭 또한 사용자의 명령어 발생음 보다 큰 일상 잡음 환경에서도 우수한 인식 성능을 나타내고 있다. 실험결과 제안한 시스템은 음성인식 시스템의 상용화에 가장 큰 걸림돌이 되는 미등록어 거절율에도 매우 우수한 성능을 보여주는 것으로 평가되었다.

또한, 본 논문에서는 제안한 호출 명령어 방식의 핵심어 검출 시스템을 실제 홈네트워크 시스템과 연동될 수 있는 임베디드 DSP 환경에서 구현함으로써

음성인식 기술을 상용화 할 수 있는 임베디드 시스템을 제시했다. 이러한 호출 명령어 기반 핵심어 검색 기술은 홈네트워크 시스템뿐만 아니라 지능형 로봇, 유비쿼터스 가전제품, 네비게이션 등 편리하고 안전한 UI 기술을 필요로 하는 다양한 적용분야의 단말에 직접 적용이 가능하다.

참 고 문 헌

[1] 한종욱, 이덕규, 정교일, "홈네트워크 보안기술 동향," 한국통신학회지, 제23권 제9호, pp. 113-124, 2006.

[2] 신상열, "홈네트워크 개요 및 정책 동향," 電波, 제128호, pp. 60-63, 2006.

[3] R. A. Sukkar and C-H. Lee, "Vocabulary independent discriminative utterance verification for non-keyword in subword based speech recognition," IEEE Trans. on speech and audio processing, Vol. 4, No. 6, pp. 420-429, Nov. 1996.

[4] Mazin G Rahim, Chin-Hui Lee, Biing-Hwang Juang and Wu Chou, "Discriminative Utterance Verification Using Minimum String Verification Error(MSVE) Training," ICASSP, 1996.

[5] Hoi-Rin Kim, SingHun Yi and Hang-Seop Lee, "Out-Of-Vocabulary Rejection using Phone Filler Model in Variable Vocabulary Word Recognition," ICSP, Vol. 1, pp. 337-339, 1999.

[6] J. Kim, J Lee, and S Choi, "Hybrid Confidence Measure for Domain-Specific Keyword Spotting," Proc. of IEA/AIE, 15, 736-745, 2002.

[7] L. Rabiner and Biing-Hwang Juang, "Fundamentals of speech recognition", Prentice-Hall International, Inc., 1993.

[8] 박진영, 권호민, 박정원, 김창근, 허강인, "DSP 기반의 잡음환경에 강인한 화자 독립 음성 인식기 구현," 대한전자공학회 03 신호처리소사이어티 추계학술대회, pp. 69-72, 2003. 11.

[9] 장홍, 정익주, "TMS320C32 DSP를 이용한 실시간 화자중속 음성인식 하드웨어 모듈(VR32) 구현," 한국음향학회지, 17권, 4호, pp. 14-22, 1998년 5월.

[10] Sadaoki Furui, "50 Years of Progress in Speech Recognition Technology-Where We Are, and Where We Should Go-," In Proc. ICASSP 2007 Plenary Speech, 2007.

[11] Deng, A. Acero, M. Plumpe, and X.D. Huang, "Large-Vocabulary Speech Recognition under Adverse Acoustic Environments," In Proc. ICSLP, 2000, pp. III-806-809.

[12] Texas Instruments, TMS320VC5501 Fixed-Point Digital Signal Processor (Rev. K). Nov 2008.



송 기 창

1984년 2월 광운대학교 응용전자공학과 학사
 1993년 8월 광운대학교 전산대학원 전자계산기공학과 석사
 2001년 2월~2003년 2월 광운대학교 전자통신공학과 박사수료
 1996년 6월~2001년 2월 삼성전자 정보통신본부 TDX 전전자교환기 개발
 2004년 2월~2010년 8월 파워보이스 HW개발팀 수석연구원
 2010년 9월~현재 아이디어폰 연구소장
 관심분야: 신호처리, 통신시스템, 하드웨어 시스템구현



강 철 호

1971년 3월~1975년 2월 한양대학교 전자공학과 학사
 1975년 3월~1979년 2월 서울대학교 전자공학과 석사
 1982년 3월~1988년 2월 서울대학교 전자공학과 박사
 1977년 3월~1982년 2월 국방과학연구소 연구원
 1983년 3월~2010년 4월 현재 광운대학교 전자통신공학과 교수
 관심분야: 디지털신호처리, 음성 및 화자인식, 멀티미디어공학