

# 분산형 음성인식 DSP 네트워킹 시스템을 위한 반음소 모델기반의 신뢰도를 사용한 결정규칙과 인터럽트-폴링

송기창<sup>†</sup>, 강철호<sup>\*\*</sup>

## 요 약

지능형 홈네트워크의 복잡하고 다양한 서비스를 음성인식을 이용해 편리하게 제어하기 위해서는 원거리 음성인식 및 분산 음성인식 네트워킹에 관한 방법은 필수적 요소이다. 이를 통해 가정의 어느 공간에서든 음성을 이용한 홈컨트롤이 가능해 진다. 본 논문에서는 분산형 음성인식 DSP 시스템 구성을 위하여 서버-클라이언트 형태로 구분된 DSP 확장형 모듈을 개발하였으며, 클라이언트 모듈이 전달하는 인식 결과의 신뢰도를 통합 분석하여 서버 모듈이 지능적으로 인식 결과를 판단하는 기법을 제안한다. 모의 실험 결과, 제안한 판단 기법은 기존의 다수결의 법칙이나, 선도착 우선의 법칙보다 우수한 인식 성능을 나타내었다. 또한, 선도착 클라이언트 결과가 도착한 후 무조건 일정 시간 타 클라이언트 결과를 기다려야 하는 기존의 지연(Delay) 방식의 단점을 해결하기 위하여, 인터럽트 폴링 기법을 제안한다. 제안한 인터럽트 폴링 기법은 서버 모듈이 클라이언트 모듈에게 현재 상태를 물어보고 클라이언트의 결과를 기다려야 할지 선택함으로써 불필요한 인식 지연시간을 없애고 인식율의 성능을 유지할 수 있다.

## Decision Rule using Confidence Based Anti-phone Model and Interrupt-Polling Method for Distributed Speech Recognition DSP Networking System

Ki-Chang Song<sup>†</sup>, Chul-Ho Kang<sup>\*\*</sup>

## ABSTRACT

Far-talking recognition and distributed speech recognition networking techniques are essential to control various and complex home services conveniently with voices. It is possible to control devices everywhere at home by using only voices. In this paper, we have developed the server-client DSP module for distributed speech recognition network system and proposed a new decision rule to decide intelligently whether to accept the recognition results or not by the transferred confidence rate. Simulation results show that the proposed decision rule delivers better performances than the conventional decision by majority rule or decision by first-arrival. Also, we have proposed the new interrupt-polling technique to remedy the defect of existing delay technique which always has to wait several clients' results for a few seconds. The proposed technique queries all client's status after first-arrival and decides whether to wait or not. It can remove unnecessary delay-time without any performance degradation.

**Key words:** Distributed Speech Recognition(분산 음성 인식), DSP Networking System, Confidence rate(신뢰도), Interrupt-Polling

※ 교신저자(Corresponding Author) : 송기창, 주소 : 서울특별시 송파구 신천동 17번지 잠실파크리오 아파트 115동 3501호(138-240), 전화 : (070)8667-1767, FAX : (032)623-8289, E-mail : kcsong78@hanmail.net  
접수일 : 2010년 2월 3일, 수정일 : 2010년 3월 16일  
완료일 : 2010년 3월 24일

<sup>†</sup> 광운대학교 전자통신공학과

<sup>\*\*</sup> 광운대학교 전자통신공학과 교수

(E-mail : chkang5136@kw.ac.kr)

※ 본 논문은 2008년도 광운대학교 연구년에 의하여 연구되었음.

### 1. 서론

최근 급속하게 발전하는 정보기술과 유비쿼터스 환경에서 이동성을 요구하는 사용자 인터페이스 기술의 개발에 의해 음성인식 분야는 많은 연구결과와 상용화 성과를 내고 있다. 특히, 지능형 홈네트워크 분야에서는 기존의 근거리 음성인식이 아니라 화자와 마이크 입력단의 거리가 3-5m 이상인 원거리 연속음성인식 시스템이 상용화되고 있는 실정이다. 이러한 원거리 음성인식 기술은 기존의 화면터치가 유일한 인터페이스이던 월패드(wall pad)의 제어를 원거리뿐 아니라 이동중에도 손쉽게 해줌으로써 가장 편리한 의사소통 도구인 음성을 사용하여 가정 내 조명, 난방, 가스 및 가전제품을 제어할 수 있도록 지원해 준다. 이러한 원거리 음성인식 확장형 모듈을 집안 곳곳에 설치하면 굳이 월패드가 있는 거실뿐 아니라 언제 어디서나 홈디바이스를 제어할 수 있는 유비쿼터스 환경을 구축할 수 있다[1-2].

본 논문에서는 유비쿼터스 홈네트워크를 실현하기 위하여 분산 음성인식 DSP 네트워크 시스템을 제안한다. 또한 잠음 환경에 취약한 음성인식 시스템의 단점을 본 논문에서 제안한 방법에 의해 극복하고자 한다. 이를 위해 서버-클라이언트 형태의 분산 구조의 네트워크를 구성하여 클라이언트로부터 전달되는 음성명령을 서버 DSP 모듈이 제안한 결정규칙에 의해 월패드로 전달하는 모델을 제시하고자 한다. 또한, 임베디드 DSP 모듈의 연산처리 속도 특성상

발생되는 불필요한 인식 지연시간을 인식율의 하락 없이 해결하도록 설계한다.

본 논문의 구성은 2장에서 분산형 음성인식 DSP 네트워크 시스템의 구성에 관해 설명하고, 3장에서는 반음소 모델 기반의 신뢰도를 적용한 인식결과에 따른 서버 모듈의 결정규칙과 인터럽트 폴링기법을 제안한다. 그리고 4장에서는 모의실험 및 결과를 검토하고, 5장에서 결론 및 고찰에 대해 각각 기술한다.

### 2. 제안한 분산형 음성인식 DSP 네트워크 시스템

원거리 음성인식을 위해서는 기본적으로 마이크 입력단의 감도가 우수해야 한다. 본 논문에서 구성된 임베디드 DSP 모듈은 -38dB의 감도를 갖는 무지향성 마이크를 사용하였다. 이러한 고감도 마이크는 화자의 발성을 뿐 아니라 주변 잡음까지도 크게 입력되어 홈네트워크용 월패드의 설치 위치에 따라 음성인식율이 크게 좌우될 수밖에 없다. 실제 가정내 환경에서 일상적으로 월패드의 위치는 TV 설치 위치와 종종 겹치게 되는데, 이런 경우, 월패드 내장형 음성인식 DSP 모듈의 음성 입력단은 화자의 발성음보다 훨씬 큰 진폭을 가지는 TV 노이즈에 의해 음성인식율이 급격히 저하된다[3-5].

본 논문에서는 이러한 음성 인식율의 저하 방지 및 가정내 어느 곳에서나 음성인식을 위해 그림 1과 같은 서버-클라이언트 형태의 분산형 음성인식 DSP

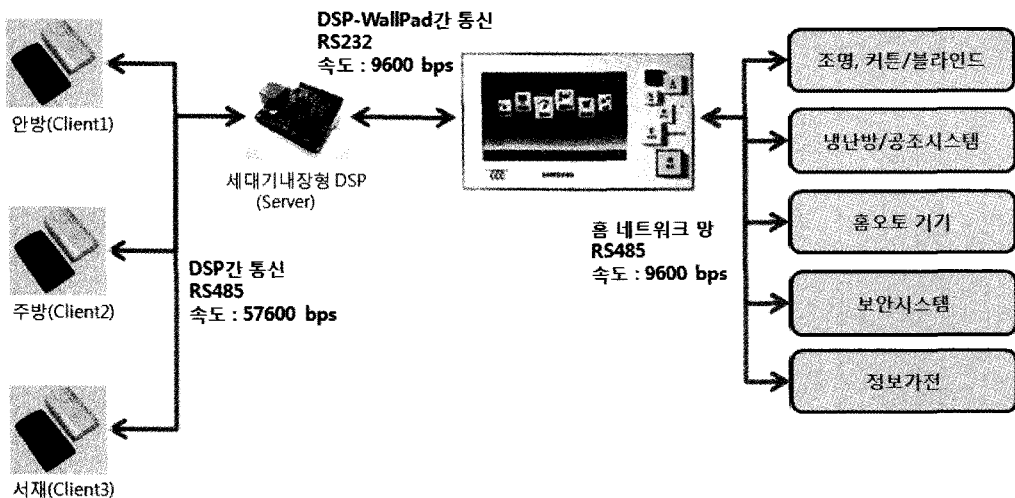


그림 1. 분산형 음성인식 DSP 네트워크 시스템의 구성도

네트워크 시스템을 제안한다. 월패드 내장형 DSP 모듈을 서버로 두고 확장형 DSP 독립 인식 모듈을 집안 곳곳에 설치하여 DSP 모듈간 RS485 통신을 통해 분산형 음성인식 네트워크를 구성한다.

각 클라이언트 DSP 모듈은 핵심어 검출 시스템으로 구성되어 독립적으로 실시간 입력되는 음성에 대해 인식을 수행하고, 인식 결과를 세대기 내장형 DSP 서버로 전달한다. 서버는 도착한 인식 결과를 결정 규칙에 따라 세대기에 전달하여 홈 네트워크망을 통해 홈 디바이스를 제어하게 된다. 제안한 분산형 음성인식 DSP 네트워킹 시스템은 잡음에 취약한 음성 인식 시스템의 단점을 보완하여 거리에 상관없이 집안 내 어느 곳에서나 세대기를 제어할 수 있다.

그림 1에서의 client 1,2,3와 세대기에 내장되는 DSP서버는 본 논문에서 제안하는 분산형 음성인식 네트워킹을 구현하고 DSP서버에 의해 그 결과를 세대기(Wall Pad)에 전달하며 세대기의 제어하에 홈네트워크 기기들이 제어 된다.

### 3. 반음소 모델기반의 신뢰도를 사용한 결정 규칙과 인터럽트-폴링 기법

제안한 음성인식 DSP 네트워킹 시스템의 서버는 클라이언트 DSP 모듈로부터 전달된 인식 결과들을 판별하여 홈네트워크 세대기로 인식결과를 전달해야 한다. 이를 결정하기 위해서 기존의 시스템은 선도착 우선 결정이나 일정 기간 시간지연을 두고 도착한 결과들의 인식 결과를 토대로 한 결정 기법을 사용하였다. 선도착 우선 결정은 가장 먼저 인식된 결과를 세대기로 전달함으로써 음성인식 처리 속도면에서 가장 우수한 장점을 가지고 있지만, 오인식율이 존재하는 음성인식 시스템의 단점을 보완하지 못한다. 따라서, 선도착한 결과가 있으면 1초-2초간 시간지연을 두어 뒤에 도착하는 인식 결과를 기다린 후, 다수결의 원칙에 따라 인식 결과를 결정하는 기법이 주로 사용되어 왔다. 그러나 지연시간 설정에 따라 선도착 우선 결정 규칙에서는 발생하지 않은 오인식이 오히려 발생하는 경우가 있기 때문에, 인식 시간은 오래 걸리면서 인식율은 크게 향상되지 않는다. 따라서 본 논문에서는 새로운 반음소 모델 기반 신뢰도를 사용한 인식 결과에 의한 서버모듈의 결정규칙을 제안한다.

### 3.1 제안한 반음소 모델기반의 신뢰도를 이용한 결정 규칙

반음소 모델을 이용한 거절 기법은 인식대상이 변화되거나 추가되는 가변어휘 연속음성 인식 시스템의 미등록어 거절기법으로 적합한 것으로 잘 알려져 있다. 하지만, 이러한 거절 기법은 미리 만들어진 음소/반음소 모델과 발성환경의 변화에 따른 불일치 조건(Mis-match condition)을 해결할 수 없다는 큰 단점을 지니고 있다. 즉, 잡음이 없는 환경에서의 음소 단위의 null hypothesis와 alternate hypothesis의 비율을 적용한 음소단위의 신뢰도와 음소단위의 신뢰도를 적용한 단어단위의 신뢰도는 등록어와 미등록어 사이의 분별력을 잘 표현해 주지만, 잡음 환경에서는 이러한 분별력이 상당히 떨어지는 경향을 보인다[6-10].

따라서, 본 논문에서는 이를 해결하기 위하여 다음과 같은 방법으로 새로운 반음소 모델을 이용한 신뢰도를 계산한다. 식 (1)은 신뢰도를 측정하기 위한 음소 대 반음소 우도비(LLR: Log Likelihood Ratio)계산 기법이다.

$$LLR = \frac{1}{N} \sum_{t=0}^N \log(p(O_t | \lambda_k) - p(O_t | \lambda_{k-anti})) \quad (1)$$

여기서,  $N$ 은 자동분할된 프레임 길이,  $O_t$ 는 관측열,  $\lambda_k$ 는 음소 모델이며,  $\lambda_{k-anti}$ 는 제안한 새로운 반음소 모델이다.

즉, 기존의 반음소 모델을 이용한 신뢰도 계산이 자신 외의 모든 음소의 평균을 내거나, 유사음소 집합을 미리 메모리에 저장하고 계산하는 것에 비하면 본 논문에서 제안하는 기법은 자신의 음소 모델만을 사용하여 음소 대 반음소 우도비를 다음과 같이 계산한다.

HMM 모델의 상태천이 확률,  $a_{ij}$ 를  $\hat{a}_{ij} = 1 - a_{ij}$ 로 대체하여 모델의 상태 천이를 자동적으로 자신의 상태 천이와 반대로 설정함으로써 반음소의 특성을 지닐 수 있다.

또한, 세그멘테이션된 프레임  $x_t (N_1 \leq t < N_T)$ 의 가우시안 모델에 대한 관측 확률 또한, 다음과 같이 프레임을 역으로 취함으로써 음소모델이 가지는 특성에 반하는 반음소의 특성을 살릴 수 있게 된다.

$$\hat{b}_j(o_t) = \frac{1}{(2\pi)^{0.5d} |\Sigma|^{0.5}} \exp[-0.5(x_{N-t} - \mu)^T \Sigma^{-1} (x_{N-t} - \mu)] \quad (2)$$

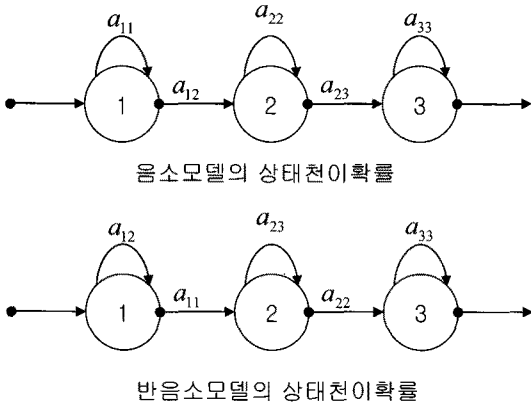


그림 2. 제안한 반응소 모델의 상태전이 확률

여기서,  $\hat{b}_j(o_t)$ 는 관측열  $o_t$ 의 심볼 관측확률,  $x_{N-t}$ 는 역으로 취해진 프레임의 MFCC,  $\mu$ 는 평균,  $\Sigma$ 는 공분산 행렬이다.

이와 같은 방법으로 신뢰도를 측정하면 음소 대 반응소의 우도비는 인식대상이 변화하거나 음성의 품질이 잡음 등에 의하여 왜곡되었을 경우에도 적응적으로 신뢰도를 측정할 수 있다. 제안한 반응소 모델의 생성 방법 및 신뢰도 계산 단계는 다음과 같다.

**step 1 :** 상태 천이 확률 대치

$$\hat{a}_{ij} = 1 - a_{ij} \tag{3}$$

**step 2 :** 심볼 관측 확률 대치

$$\hat{b}_j(O_t) = \frac{1}{(2\pi)^{0.5d} |\Sigma|^{0.5}} \exp[-0.5(x_{N-t} - \mu)^t \Sigma^{-1} (x_{N-t} - \mu)] \tag{4}$$

여기서,  $\lambda_k$ 의 반응소 모델은  $\lambda_{k-anti} = \{\pi, \hat{a}_{ij}, \hat{b}_j(o_t)\}$ 로 대치된다.

**step 3 :** 신뢰도 계산

$$\hat{\delta}_i(j) = \max_{1 \leq i \leq N} [\hat{\delta}_{i-1} \hat{a}_{ij} \hat{b}_j(o_t)] \tag{5}$$

신뢰도,  $LLR \quad \omega = \frac{1}{N} \sum_{t=0}^N \log(p(O_t | \lambda_k) - p(O_t | \lambda_{k-anti}))$  (6)

본 논문에서 제안하는 새로운 반응소 모델을 이용한 신뢰도 계산법은 학습시 환경과 인식 환경의 차이에서 기인하는 채널 불일치(Mis-match condition) 조건 발생시 큰 오류가 발생하는 기존의 신뢰도 계산 기법에 비해 잡음에 강인하며 특별히 반응소 모델을 따로 구축할 필요가 없다. 서버 모듈은 각 클라이언트 모듈로부터 전달 받은 신뢰도,  $\omega$ 를 (7)식과 같이

인식 인덱스 별로 취합, 평균을 구한 후, 최대의 신뢰도를 갖는 인덱스를 최종 인식 결과로 결정한다.

$$I = \arg \max_{1 \leq i \leq N} [ave(w_i)] \tag{7}$$

여기서,  $I$ 는 최종 인식 결과의 인덱스,  $ave(w_i)$ 는  $i$ 번째 클러스터 인식결과와 평균이다.

### 3.2 제안한 인터럽트-폴링 기법

기존의 star형 분산 음성인식 시스템의 구조는 그림 3과 같이 음성인식 클라이언트 모듈이 각각 음성인식을 수행하다 인식에 성공하면 서버 모듈로 인터럽트(Interrupt)를 보내게 되고, 서버 모듈은 최초 인터럽트 신호 후 다른 클라이언트 모듈에서 음성인식이 진행되고 있기 때문에 일정시간 기다리게 된다 [11]. 일정시간 후 각 모듈의 신뢰도를 이용하여 제안한 결정 규칙에 따라 월패드로 전송함으로써 홈제어를 가능케 한다.

기존의 star형 음성인식 네트워크에서는 최초 인터럽트 신호가 들어온 후 항상 1-2초의 일정시간 대기해야 한다는 단점을 가진다. 임베디드 DSP 환경에서 화자가 명령을 발성한 후 단말기의 응답을 기다리기까지의 평균 응답시간이 1.0-1.2초인 것을 감안할 때, 1-2초의 응답 지연시간은 전체 인식 시스템의 성능이 저하되는 느낌을 받기에 충분하다.

따라서, 본 논문에서는 그림 4와 같은 인터럽트-폴링 방식을 제안한다.

제안한 인터럽트-폴링 방식의 수행 단계는 다음과 같다.

**step 1 :** 클라이언트 모듈은 각각 음성인식을 처리하다가 인식이 성공하면 서버 모듈에 인터럽트를 보내어 결과를 알린다.

**step 2 :** 서버 모듈은 최초 인터럽트 신호를 받으면 각각의 클라이언트 모듈에 폴링, 현재 처리되고

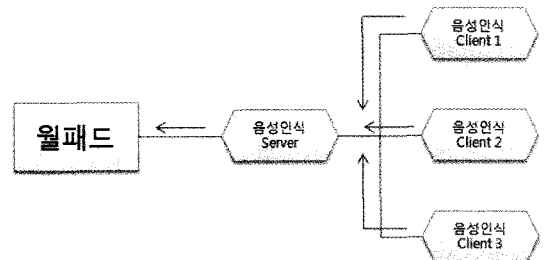


그림 3. 기존의 인터럽트-딜레이 방식의 모식도

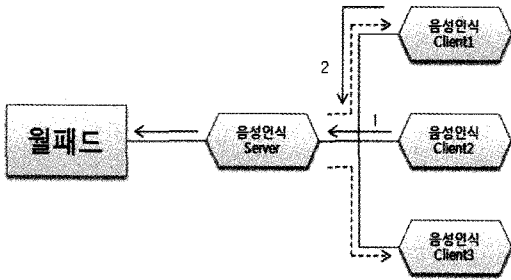


그림 4. 제안한 인터럽트-폴링 방식의 모식도

있는 상태를 묻는다.

*step 3*: 처리되는 것이 없으면 바로 월패드로 정보를 전달하고, 현재 진행중인 모듈이 있다면, 처리될 때 까지 기다린 후 신뢰도를 이용한 결정 규칙에 따라 인식 결과를 월패드로 전송한다.

제안한 방식은 불필요한 딜레이를 없애고, 응답 속도를 개별 모듈의 수준으로 향상시킬 수 있다. 또한, 정확히 인식중인 클라이언트 모듈의 ID를 알 수 있기 때문에 인식중인 모듈의 모든 결과를 통합 분석할 수 있는 장점을 지닌다.

#### 4. 모의 실험 및 결과

##### 4.1 실험 환경

모의 실험에 사용한 tri-phone 모델 훈련을 위하여 6종류(ETRI의 PBW445 DB, POW3848 DB, 국어

공학연구소의 PBW452DB, 4연숫자음, 단독숫자음, 고빈도2000어휘DB)와 홈네트워크용 545 DB)를 사용하였다. 학습에 참여한 음성DB는 문맥 종속적 모델인 2-mixture tri-phone으로써 묵음(silence)모델을 포함하여 총 18,349개이다. 임베디드 DSP 환경에서 실시간 인식 처리할 수 있도록, 8kHz down-sampling 과정을 거친 후, 30msec 구간을 1 프레임으로 분석하여 Energy, 12MFCC(Mel-scaled Frequency Cepstral Coefficient), Delta-MFCC의 26차 특징 파라미터를 추출하였고, 추출된 파라미터를 이용하여 EM(Expectation-Maximization) 알고리즘에 의해 tri-phone 모델링하였다. 각 tri-phone은 3상태의 left-to-right 전이를 가지는 상태 모델로 훈련하였다. 각 상태는 계산량 대비 우수한 성능을 나타낸 2-mixture 혼합 모델로 구성하였다.

평가용 Test DB는 실제 가정 환경에서 거실 4곳에 각각 마이크 및 노트북을 설치하고, 3종류의 TV 노이즈 레벨 (40, 50, 60dB) 환경에서 동시에 발생시켜 구축하였다. Test DB 구축에 참여한 화자는 20-40대 연령대별 남녀 화자 20명씩 총 120명이 녹음하였다.

##### 4.2 실험 결과 및 검토

표 1, 2, 3은 세가지 종류의 잡음 환경에서 DSP 단독모듈, 기존 star형 인터럽트-딜레이 네트워크 시스템 및 제안한 분산형 음성인식 DSP 네트워크 시스템

표 1. 40dB 잡음환경에서의 기존 및 제안한 시스템의 음성 인식률

평가항목	DSP 단독모듈		기존 star형 인터럽트-딜레이 기법		제안한 신뢰도를 이용한 결정규칙		제안한 인터럽트-폴링 기법	
	남	여	남	여	남	여	남	여
1. 음성인식률(40dB잡음환경)	95%	93%	97%	97%	98%	99%	98%	98%
2. 오인식률	3%	3%	2%	3%	1%	1%	1%	1%
3. 응답처리속도	1.0 sec		2.5 sec		2.5 sec		1.5 sec	

표 2. 50dB 잡음환경에서의 기존 및 제안한 시스템의 음성 인식률

평가항목	DSP 단독모듈		기존 star형 인터럽트-딜레이 기법		제안한 신뢰도를 이용한 결정규칙		제안한 인터럽트-폴링 기법	
	남	여	남	여	남	여	남	여
1. 음성인식률(50dB잡음환경)	90%	89%	94%	93%	95%	94%	95%	93%
2. 오인식률	3%	2%	3%	3%	1%	1%	1%	1%
3. 응답처리속도	1.2 sec		2.5 sec		2.5 sec		1.6 sec	

표 3. 60dB 잡음환경에서의 기존 및 제안한 시스템의 음성 인식율

평가항목	DSP 단독모듈		기존 star형 인터럽트-딜레이 기법		제안한 신뢰도를 이용한 결정규칙		제안한 인터럽트-폴링 기법	
	남	여	남	여	남	여	남	여
1. 음성인식율(60dB잡음환경)	80%	75%	90%	89%	92%	93%	92%	92%
2. 오인식율	2%	2%	3%	4%	1%	1%	1%	1%
3. 응답처리속도	1.5 sec		3.0 sec		3.0 sec		1.7 sec	

템의 성능 평가표이다. DSP 단독모듈과 기존의 star형의 경우, 잡음이 증가함에 따라 인식율이 급격히 하락하지만, 제안한 시스템의 경우, 인식율 하락이 둔화됨을 알 수 있다. 특히 오인식율의 경우, 제안한 2가지 방식 모두 잡음환경에 관계없이 1%의 안정된 수준을 유지하고 있다.

그리고 제안한 신뢰도를 이용한 결정규칙의 네트워킹 시스템은, 잡음원으로부터 상대적으로 가장 멀리 위치하는 클라이언트 DSP 모듈들에 의한 인식 신뢰도(음성인식율, 오인식율)가 모든 잡음환경에서 우수한 성능을 보여준다. 또한, 제안한 인터럽트-폴링 기법은 DSP 단독모듈에 비해 0.2-0.5초 차이의 빠른 응답처리속도를 보이고 있으며, 제안한 신뢰도를 이용한 결정규칙의 음성인식율과 오인식율 결과와 비교하더라도 거의 유사한 수준을 유지하고 있음을 확인할 수 있다. 또한, 응답처리속도는 제안한 인터럽트-폴링 기법이 신뢰도를 이용한 결정규칙에 비해 대략 2배 정도 빠르다는 것을 알 수 있다.

### 5. 결론 및 고찰

본 논문은 홈네트워크용 음성인식 DSP 모듈을 집안내 곳곳에 분산 설치하여 서버-클라이언트 형태로 네트워킹 시스템을 구축함으로써 유비쿼터스 환경 모델을 제시하였다. 사용자는 가정내 어느 곳에서라도 음성을 이용한 편리한 홈다바이스 제어가 가능할 뿐 아니라, 사용자의 발성음 대 잡음 비율(SNR : Signal to Noise Ratio)을 극대화시킴으로써 잡음환경에 취약한 음성인식 시스템의 단점을 극복할 수 있다.

본 논문에서 제시하고 있는 시스템은 학습시 채널과 인식시 채널 변화에서 오는 채널 불일치(Channel Mis-match) 조건을 새로운 반응소 모델을 이용하여 제거하고, 이로 인한 신뢰도를 이용한 서버의 음성인

식 결과에 따른 결정규칙은 유사한 음향학적 특성 때문에 혼동될 수 있는 인식 오류율을 획기적으로 감소시킬 수 있다. 또한, 최초로 인터럽트 신호가 도착한 후 각 클라이언트의 상태를 확인함으로써 불필요한 대기 지연 시간을 감소시키고 인식율의 하락을 막을 수 있다. 따라서 제안한 음성인식 네트워킹 시스템은 화면 터치에 의존하고 이동성이 제한된 홈네트워크 시스템에 많은 변화를 가져오리라고 기대된다.

### 참고 문헌

- [1] 강재구, 김성완, 임승린, “홈 네트워크(Home Network) 트렌드와 모델,” 韓國컴퓨터情報學會誌, 제16권 제2호, pp. 27-32, 2008.
- [2] 이호용, 정희석, “지능형 홈네트워크 시스템을 위한 가변어휘 연속음성인식시스템에 관한 연구,” 韓國ITS學會 論文誌, 제7권 제2호, pp. 37-42, 2008.
- [3] Lin, Q., Che, C., Yuk, D.-S., Jin, L., De Vries, B., Pearson, J and Flanagan, J. “Robust Distant Talking Speech Recognition,” In Proc. ICASSP, v.1, pp. 21-24, 1996.
- [4] 권석봉, 지미경, 김희린, 이용주, “원거리 음성인식을 위한 MLLR적용기법 적용,” 대한음성학회 말소리, 제53호, pp. 119-127, 2005.
- [5] Nishiura, T., Hirano, Y., Denda, Y and Nakayama, M. “A study of suitable reverberation criteria for distant-talking speech recognition with room acoustic parameters,” International Congress on Acoustics, pp.CAS-03-036, 2007.
- [6] Hoi-Rin Kim, SiongHun Yi and Hang-Seop Lee, “Out-of-Vocabulary Rejection using Phone Filler Model in Variable Vocabulary

- Word Recognition,” ICSP, v.1, pp.337-339, 1999.
- [7] 김우성, 구명완, “반음소 모델링을 이용한 거절 기능에 대한 연구,” 한국음향학회지, 제18권, 제3호, pp. 3-9, 1999.
- [8] Sunil K. Gupta and Frank K. Soong, “Improved Utterance Rejection Using Length Dependent Thresholds”, ICSLP, pp. 795-798, 1998.
- [9] Li Jiang and Xuedong Huang, “Vocabulary-Independent Word Confidence Measure Using Subword Features”, ICSLP, pp. 401-404, 1998.
- [10] M. Rahim, et al., “Discriminative utterance verification using minimum string verification error (MSVE) training,” Proc. IEEE- ICASSP, pp. 3585-3588, 1996.
- [11] 지미경, 김회린, “다채널 마이크 환경에서 Naive Bayesian Network의 Decision에 의한 음성인식 성능향상,” 대한음성학회 05 추계 학술대회 발표논문집, pp. 151-154, 2005 Nov.



송기창

1984년 2월 광운대학교 응용전자공학과 학사  
 1993년 8월 광운대학교 전산대학원 전자계산기공학과 석사  
 2001년 2월~2003년 2월 광운대학교 전자통신공학과 박사수료  
 1988년 5월~1996년 5월 한화정보통신(구 동양전자통신) TDX개발단 TDX 전전자교환기 개발  
 1996년 6월~2001년 2월 삼성전자 정보통신본부 TDX 전전자교환기 개발  
 2003년 3월~2007년 2월 동서대학교, 동원대학, 신홍대학 정보통신과 겸임교수  
 2004년~현재 파워보이스 HW개발팀 수석연구원  
 관심분야: 신호처리, 통신시스템, 하드웨어 시스템구현



강철호

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