

한국어 유아 음성인식을 위한 수정된 Mel 주파수 캡스트럼 Modified Mel Frequency Cepstral Coefficient for Korean Children's Speech Recognition

유재권*, 이경미**
(주)파수닷컴*, 덕성여자대학교 컴퓨터학과**

Jae-Kwon Yoo(jkyoo@fasoo.com)*, Kyoung-Mi Lee(kmlee@duksung.ac.kr)**

요약

본 논문에서는 한국어에서 유아 대상의 음성인식 향상을 위한 새로운 특징추출 알고리즘을 제안한다. 제안하는 특징추출 알고리즘은 세 가지 방법을 통합한 기법이다. 첫째 성도의 길이가 성인에 비해 짧은 유아의 음향적 특징을 보완하기 위한 방법으로 성도정규화 방법을 사용한다. 둘째 성인의 음성과 비교했을 때 높은 스펙트럼 영역에 집중되어 있는 유아의 음향적 특징을 보완하기 위해 균일한 대역폭을 사용하는 방법이다. 마지막으로 실시간 환경에서의 잡음에 강건한 음성인식기 개발을 위해 스무딩 필터를 사용하여 보완하는 방법이다. 세 가지 방법을 통해 제안하는 특징추출 기법은 실험을 통해 유아의 음성인식 성능 향상에 도움을 준다는 것을 확인했다.

■ 중심어 : | 음성인식 | 음성 데이터베이스 | 음성 인터페이스 | MFCC |

Abstract

This paper proposes a new feature extraction algorithm to improve children's speech recognition in Korean. The proposed feature extraction algorithm combines three methods. The first method is on the vocal tract length normalization to compensate acoustic features because the vocal tract length in children is shorter than in adults. The second method is to use the uniform bandwidth because children's voice is centered on high spectral regions. Finally, the proposed algorithm uses a smoothing filter for a robust speech recognizer in real environments. This paper shows the new feature extraction algorithm improves the children's speech recognition performance.

■ keyword : | Speech Recognition | Speech Database | Speech Interface | MFCC |

I. 서론

음성은 사람들 사이에서 가장 자연스럽게 편리한 정보 교환의 수단으로서, 음성을 이용한 인터페이스는 사람과 컴퓨터와의 인터페이스(HCI : Human-Computer Interfac) 연구 분야에서 핵심 기술 중 하나이다. 음성

인터페이스는 그 편의성 때문에 많은 어플리케이션에 적용하려는 시도가 이루어지고 있다[1][2].

특히, 유아 대상의 교육용 프로그램에서 음성 인터페이스의 사용은 유아에게 다양한 형태의 경험을 제공하고 언어 발달 능력을 향상시킬 수 있는 장점을 가지고 있다. 그러나 현재의 음성 인터페이스는 성인 음성 데

이터의 접근 용이성 때문에 성인 대상으로 구축되어 있다. 이전 연구에 따르면[3], 성인 음성 인터페이스를 유아 음성에 적용하였을 때, 유아 음성의 특징을 반영하지 못하여 상당히 낮은 성능을 확인했다. 따라서 유아의 음성에 맞도록 성능을 향상시킬 수 있는 연구가 필요하다.

음성 인터페이스를 개발하기 위해서 가장 중요한 기술은 음성 인식이다. 음성인식 기술은 음성데이터에서 좋은 특징을 추출하는 것이 중요한데, 음성인식을 위해 주로 사용되는 특징은 LPC 캡스트럼, PLP 캡스트럼, Mel 주파수 캡스트럼(MFCC : Mel Frequency Cepstral Coefficient), 필터뱅크 에너지 등이 있다. 이 중 MFCC는 인간의 청각 주파수 특성을 반영한 것으로 일반적으로 음성인식 수행에 가장 좋은 효과를 주는 방식으로 많이 쓰이고 있는 특징추출 기법이다.

유아의 음성은 성인과 다른 음성 특징을 갖고 있다. 우선, 유아의 음성은 성인의 음성보다 상대적으로 큰 스펙트럼과 일시적인 변이를 가지고 있다[4]. 이런 변이는 보통 성도 길이에 따른 해부학적인 차이 때문에 일어나게 된다. 이런 이유로 성도길이 정규화(VTLN : Vocal Tract Length Normalization)를 MFCC에 적용하면 유아의 음성인식 향상에 중요한 역할을 할 수 있다[5-7]. 또한, 유아의 음성은 성인에 비해 상대적으로 높은 스펙트럼 영역에 집중되어 있다. 이런 높은 스펙트럼에 집중된 음성을 잘 표현하기 위해서 MFCC에서 사용되는 필터의 수, 필터의 모양, 대역폭의 필터 뱅크를 변형시킬 수 있는데[8], 균일한 필터 뱅크를 적용하면 유아의 음성인식 향상에 큰 영향을 미칠 수 있을 것이다[9].

본 논문에서는 한국어에서 유아의 음성인식 향상을 위해서 특정 MFCC의 필터를 변형하고 VTLN을 적용하는 방법을 제안한다. 또한, 실시간 환경에서의 잡음에 강건한 음성인식기 개발을 위해 스무딩 필터를 사용하여 보완하였다. 제안된 방법은 한국 유아음성 데이터베이스를 이용하여 실험이 이루어졌으며[10], 기존의 MFCC에 비해 제안하는 방법이 한국어 유아 음성인식 향상에 도움을 준다는 것을 확인한다.

II. 제안하는 알고리즘

본 장에서 제안하는 SUV(Smoothing Uniform VTLN)-MFCC 알고리즘은 다음과 같은 유아 음성의 특징을 이용한다: 유아의 성도 길이는 성인에 비해 짧고, 유아의 음성은 성인에 비해 높은 스펙트럼 영역에 집중되어 있다. 이러한 점을 보완해 주어야만 유아 음성의 특징을 적절하게 추출할 수 있다. 또한 실제 환경에서 잡음에 강건한 음성인식기 개발을 위해 평활화 필터를 사용한다.

1. 성도길이 정규화(VTLN)

VTLN은 화자 상호간의 차이를 줄이기 위해 음성 신호의 파라미터 성분을 정규화하는 기술이다. VTLN의 가장 중요한 개념은 인식과정에서 음향학적 벡터들에서 화자 의존적 가변성을 제거하기 위해 각 화자를 위한 음향학적 벡터들의 주파수 축을 정규화하는 것이다. 일반적인 발성에 대하여 스펙트럼 포먼트 정점들의 위치는 성도 길이에 반비례한다. 유아의 성도 길이는 성인에 비해 짧은데, 이는 유아가 발성되는 소리에서의 스펙트럼 주파수가 높다는 것을 의미한다. 이런 부분을 보완하기 위해 VTLN을 사용하여 주파수 스펙트럼을 이동하여 보상하는 주파수 와핑(warping)을 수행한다.

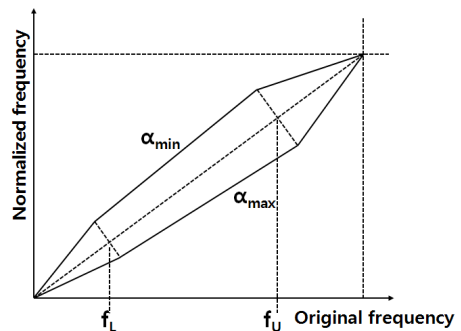


그림 1. 단순 선형와핑을 사용한 성도길이 정규화

VTLN은 화자에 따른 성도 길이 변화에 영향을 피하기 위해 주파수 축을 확대하거나 축소하여 정규화하는 방법으로 주파수 축의 척도를 변환하는 와핑 함수를 사

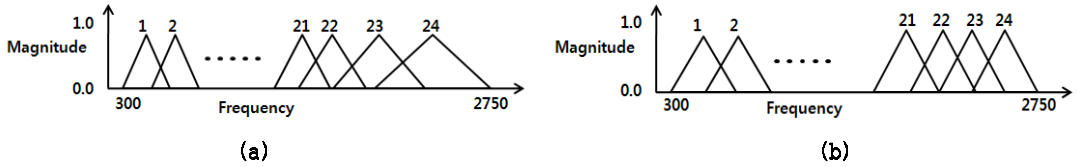


그림 2. (a) Constant-Q필터를 사용한 일반적인 MFCC와 (b) 균일한 필터를 사용한 변형된 MFCC

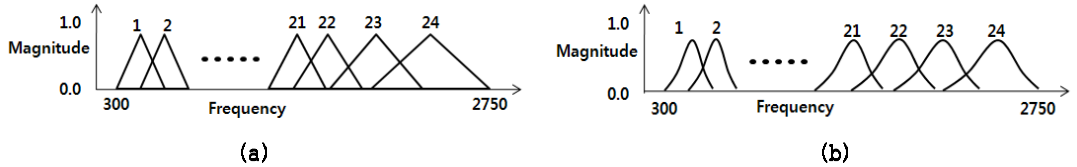


그림 3. (a) MEL 필터를 사용한 일반적인 MFCC와 (b) 평활화 필터를 사용한 변형된 MFCC

용한다. [그림 1]과 같은 단순한 선형 와핑은 f_L 과 f_U 의 변수에 의해 넓은 주파수 대역 내에서 특정 주파수를 기준하여 주파수 대역을 나누게 된다. 본 실험에서는 f_L 은 300Hz, f_U 은 2,750Hz가 사용되었다.

또한, 주파수 와핑 요소는 일반적으로 α 로 표기할 수 있는데, 단순한 선형 와핑 요소 α 값은 생리학상의 언급을 기반으로 0.8~1.22의 범위를 갖는 것이 일반적인 가정이다. $\alpha < 1.0$ 이면 변환된 주파수 축을 축소하고, $\alpha > 1.0$ 이면 축을 확대한다. 유아의 경우 짧은 성도 길이로 스펙트럼의 주파수가 높기 때문에 이를 보완하기 위해 와핑 요소는 $\alpha < 1.0$ 을 사용하여 주파수 축을 축소하는 것이 필요하다. 본 논문에서는 $\alpha < 1.0$ 에 포함된 값인 0.8을 사용했다.

2. 균일(Uniform) 필터

유아의 음성인식 향상을 위해 MFCC에서 사용하는 필터의 대역폭을 바꿀 수 있다. 본 논문에서는 f_L 은 300Hz, f_U 은 2,750Hz인 주파수 대역에 대해 24개의 필터를 사용하였다. [그림 2]는 24개의 필터에 대해 일반적인 MFCC에서 사용하는 대역폭과 제안하는 MFCC의 대역폭을 보여주고 있다. [그림 2](a)는 일반적인 MFCC에서 사용하는 constant-Q 필터 뱅크로 높은 주파수에서 더 넓은 대역폭을 갖는다. 반면, [그림 2](b)의 대역폭은 낮은 주파수에서 높은 주파수까지 동일한 대역폭을 사용하는 균일한 필터 뱅크를 사용한다.

유아의 음성은 성인의 음성과 비교했을 때 상대적으로

로 높은 스펙트럼 영역에 집중되어 있는 특징이 있다. 따라서 일반적인 MFCC에서 사용하는 constant-Q를 사용하게 되면 오히려 높은 주파수에서의 대역폭이 넓어지기 때문에 필터링 과정에서 정확한 음성 데이터를 표현하지 못해 손실이 생길 수 있다. 따라서 본 논문에서는 이러한 유아의 음성 특징을 보완할 수 있는 균일 필터 사용을 제안한다. 균일한 필터를 사용하게 되면 낮은 주파수에서 높은 주파수까지 동일한 대역폭을 사용하는 필터를 사용하기 때문에 높은 스펙트럼 영역에 집중되어 있는 유아의 음성 데이터를 상대적으로 조밀하게 표현할 수 있어 데이터 손실을 줄일 수 있다.

3. 평활화(Smoothing) 필터

음성인식을 실제로 사용가능하게 만들기 위해서는 주변에서 발생하는 잡음에 대한 처리를 고려하여야 한다. 이 과정에서 본래 데이터의 손실이 있을 수 있지만 불필요한 노이즈를 제거하는 과정이 필요하기 때문이다.

[그림 3]은 일반적인 MFCC에서 사용되는 삼각필터와 평활화 필터를 보여주고 있다. 단기간 음성에 대해 [그림 3](a)의 삼각필터를 이용하여 파워스펙트럼을 구한다. 이 과정에서 필터들이 겹치는 현상을 볼 수 있는데 이는 실시간으로 들어오는 음성에 대해 필터링하는 과정에서의 음성 손실을 막기 위한 것이다.

본 논문에서는 삼각 필터 모양 대신 [그림 3](b)와 같이 평활화된 필터 모양의 사용을 제안한다. 일반적인 MFCC에서는 평균적인 스펙트럼을 구한 값을 이용해

파워스펙트럼을 구하는 반면, 평활화된 필터 모양은 각각의 윈도우된 서브 프레임인 푸리에 변환(F)의 크기를 제공함으로써 계산되어진 파워스펙트럼이다.

$$G_{av}(f) = G(f) * \{F[w(t)]\}^2$$

III. 실험 결과

1. 실험 구성

본 논문에서는 한국어에서 만 3~5세 유아의 음성인식 향상을 위한 변형된 특징추출 방법을 제안한다. 제안된 방법의 유효성을 실험하기 위해, 유아가 많이 사용하는 단어 중 13개 분야에서 105개를 선정하였다[표 1][10]. 105개 단어에 대해 40명의 남녀 유아로부터 수집한 12,600개의 음성데이터를 사용하였다[10]. 음성 데이터는 16KHz 샘플링 비율, Mono 채널, 16bit로 저장되었다. 실험시 훈련 데이터와 테스트 데이터는 [표 2]와 [표 3]과 같다. 실험은 HTK(HMM Tool Kit)에 기반한 인식을 통해 진행되었다[11].

표 1. 음성 데이터[10]

개념	색	빨강, 파랑, 노랑, 초록, 분홍, 검정, 보라, 흰색, 주황
	형	동그라미, 세모, 네모, 별, 달, 하트, 삼각형, 사각형, 원
	수	일, 이, 삼, 사, 오, 육, 칠, 팔, 구, 십 하나, 둘, 셋, 넷, 다섯, 여섯, 일곱, 여덟, 아홉, 열
	시간	먼저, 나중, 지금, 아직, 내일, 어제, 오늘, 아까
위치	위, 아래, 오른쪽, 왼쪽, 앞, 뒤, 옆, 안, 밑	
	보기	확대, 축소, 회전
지시어	실행	예, 아니오, 시작, 다시, 그만, 처음으로
	동물	여우, 사자, 코끼리, 토끼, 기린, 호랑이, 말, 오리, 돼지
단어	탈것	기차, 자동차, 배, 오토바이, 버스, 비행기, 트럭
	직업	간호사, 경찰관, 소방관, 우체부, 의사, 선생님
	옷	장갑, 양말, 치마, 외투, 잠바, 잠옷, 바지
	장소	병원, 유치원, 은행, 놀이터, 동물원, 학교, 수영장
	장난감	로봇, 인형, 블록, 풍선, 비누방울

표 2. 훈련에 참여한 유아 분포

나이(만)	3	4	5	합계
남	3	4	3	10
여	4	3	3	10
합계	7	7	6	20

표 3. 테스트에 참여한 유아 분포

나이(만)	3	4	5	합계
남	3	3	4	10
여	3	3	4	10
합계	6	6	8	20

2. 실험 결과

본 논문에서는 제안하는 SUV-MFCC의 한국어 유아 음성인식률을 보여주기 위해 SUV-MFCC에서 사용하는 VTLN, 균일 필터, 평활화 필터가 유아 음성인식에 얼마나 기여하는지에 대한 실험을 진행하였다.

2.1 VTLN 와핑 요소에 따른 SUV-MFCC 비교 실험

우선, VTLN의 효과를 확인하기 위해 균일 필터와 평활화 필터를 사용했고, VTLN만 다양한 와핑 요소 α 값을 적용했다. 따라서, 본 실험에서는 주파수 축을 축소시키는 와핑 요소 α 값인 0.8이외에도, 아무런 영향을 미치지 않는 1.0, 반대로 축을 늘리는 1.22를 사용했다. α 값이 0.8일 경우에 비해, 1.0일 경우엔 평균 89.60%로 3.39%, 1.22일 경우엔 평균 86.61%로 6.38%의 인식률 저하를 보여주었고 있다. 이는 작은 α 값이 한국어에서 짧은 성도를 가진 유아의 높은 피치의 소리를 보완해 주는 반면, 큰 α 값은 높은 피치의 특성을 약화시켜 VTLN을 사용하지 않을 때보다 인식률이 좋지 않음을 의미한다.

또한, 단어 분야별로 VTLN을 적용하지 않은 경우와 비교할 때[그림 4], VTLN 와핑 요소 0.8의 사용은 장난감, 직업, 수, 형, 위치, 색, 탈것, 장소에 대한 단어에서 3% 이상의 높은 인식률 향상을 보여주었다. 특히, '노랑'과 '파랑', '코끼리'와 '토끼', '세모'와 '네모', '삼각형'과 '사각형' 처럼 동일 음절을 포함한 단어들의 혼동을 6~15% 이상 개선하였다. 따라서, 작은 α 값은 유아의 높은 스펙트럼을 세분화하기 때문에 높은 피치의 유아 음성과 더불어 동일 음절을 포함한 단어의 음성인식에 도움을 줄 수 있음을 확인했다.

2.2 대역폭에 따른 SUV-MFCC 비교 실험

대역폭에 따른 비교 실험을 위해, VTLN의 와핑 요소 값 0.8과 평활화 필터를 사용했고, 대역폭만 균일 필

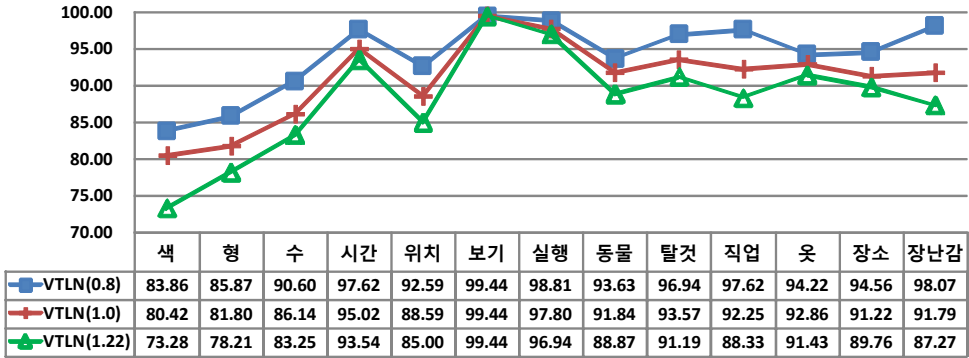


그림 4. 인식률(%) : SUV-MFCC의 다양한 VTLN 와핑 요소에 따른 인식 결과

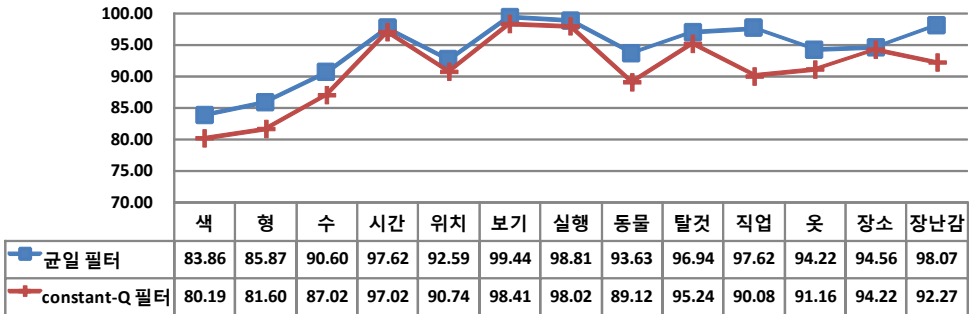


그림 5. 인식률(%) : SUV-MFCC의 대역폭에 따른 인식 결과

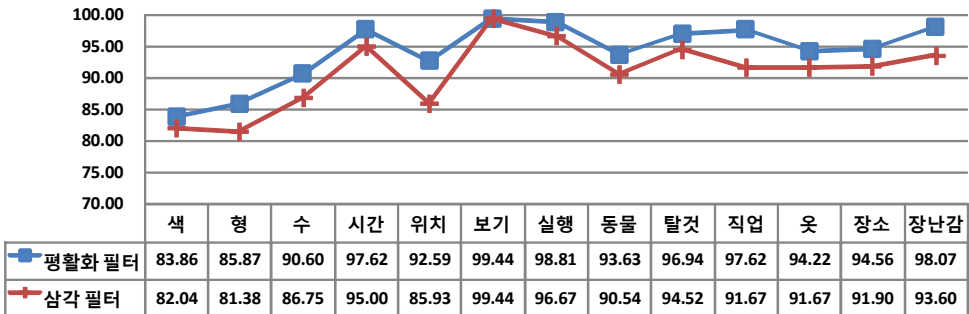


그림 6. 인식률(%) : SUV-MFCC의 필터 모양에 따른 인식 결과

터와 constant-Q 필터 대역폭을 적용했다. 균일 필터의 사용했을 때 보다 constant-Q 필터를 사용했을 경우엔 평균 89.91%로 3.07%의 인식률이 저하가 일어났다. 이는 균일 필터가 유아처럼 높은 주파수를 가진 음성의 특징을 보다 자세히 표현할 수 있음을 의미한다고 볼 수 있다.

또한, 단어별로 보았을 때[그림 5], 실험에 사용하는 대부분의 단어에서 2~8% 이상의 인식률 향상을 이루었다. 특히, 주목할 부분은 ‘검정’, ‘지금’, ‘아까’, ‘그만’, ‘코끼리’, ‘토끼’, ‘기린’, ‘블록’ 등 파열음이 속해 있는 단어들이 높은 인식률 향상을 보였다는 것이다. 파열음은 날숨을 조음기관으로 막았다가 갑자기 터트릿듯 내보

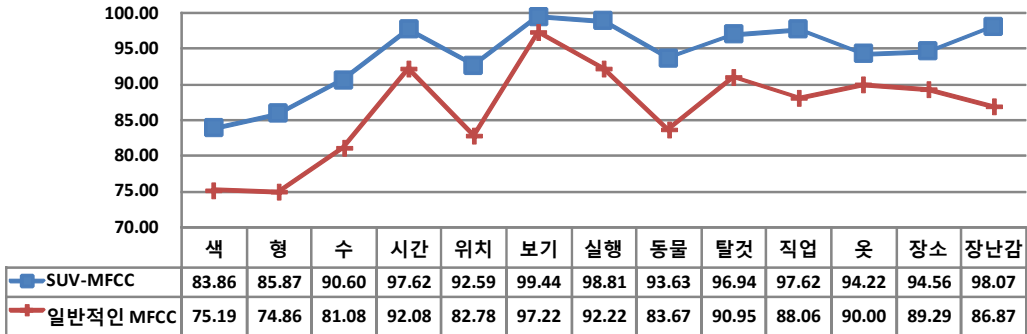


그림 7. 인식률(%) : SUV-MFCC와 일반적인 MFCC의 인식 결과

내면서 내는 소리로, 에너지는 작지만 높은 주파수 성분을 가지는 특성을 갖고 있다. 따라서 균일 대역폭을 사용함으로써 유아의 음성의 특징인 높은 주파수 대역의 정보 손실을 막을 수 있었고, 동시에 과열음에 관한 특성을 보완하여 이와 같은 높은 음성인식 향상을 구할 수 있었다.

2.3 필터 모양에 따른 SUV-MFCC 비교 실험

필터 모양에 따른 효과를 확인하기 위해, VTLN의 와핑 요소 값 0.8과 균일 필터를 사용했고, 필터 모양만 평활화 필터와 삼각 필터를 적용했다. 평활화 필터를 사용했을 때보다 삼각 필터를 사용할 경우 평균 89.50%로 약 3.49%의 인식률을 저하를 보여주었다. [그림 6]은 평활화 필터가 단어별로 고르게 잡음으로 인한 인식률 저하를 보상해 주고 있음을 보여주고 있으며, 위치와 직업은 6%이상의 인식률 향상을 이루었다.

특히, 평활화 필터를 사용하는 이유는 잡음에 의한 인식 수행 저하를 방지하기 위함이므로, 음성 환경에 따라 인식률 향상에 영향을 받을 수 있다. 본 실험에 사용된 음성 데이터베이스는 유치원의 실제 환경을 담기 위해서 잡음을 완전히 정제시키기 보다는 유치원의 비교적 조용한 별도의 방에서 음성 데이터를 수집하여 작은 잡음과 울림만이 부분적으로 포함되어 있다. 이번 실험을 통해 평활화 필터가 실제 환경에서 잡음에 강건한 음성인터페이스 개발에 도움이 될 수 있음을 확인했다.

2.4 종합

제안하는 SUV-MFCC는 VTLN, 균일 필터, 평활화 필터를 모두 사용하는데, 이를 사용하지 않는 일반적인 MFCC와 비교실험을 수행하였다. 한국어 유아 음성에 대해 SUV-MFCC의 인식률은 평균 92.99%로 일반적인 MFCC의 인식률 84.82%보다 약 8.17% 더 높은 인식률을 획득했다. 특히, [그림 7]은 장난감과 형에 대한 단어는 평균 11% 이상, 동물, 위치, 직업, 수에 대한 단어는 평균 9% 이상의 높은 인식률 향상을 보여주고 있다.

또한, 단어별로 보았을 때, ‘코끼리’가 33.33%, ‘노랑’이 28.57%, ‘삼’이 21.43%의 높은 인식률 향상을 이루었다. 또한, ‘빨강’, ‘파랑’, ‘초록’, ‘밀’, ‘간호사’, ‘경찰관’은 14%이상의 향상이 있었다. 반면, 여진희 ‘넷’은 ‘셋’과 혼동, ‘다섯’은 ‘여섯’과 혼동하는 현상을 보였다.

IV. 결론

본 논문에서는 한국어에서 유아의 음성인식 향상을 위한 특징 추출 방법을 제안했고, 다양한 실험을 수행했다. 일반적으로 사용되고 있는 MFCC를 이용하여 비교 실험을 하였고, 제안하는 알고리즘 SUV-MFCC에서 VTLN, 필터 모양, 대역폭 변환이 어떤 영향을 미치는지도 실험했다. 이 실험을 통해 다양한 방법으로 새롭게 제안된 SUV-MFCC 알고리즘이 성인과 달리 성도 길이가 짧기 때문에 높은 스펙트럼에 집중되어 있는 유아의 음성 특징을 보완하는 알고리즘으로서의 가능

성을 입증했다. 제안된 음성특징 추출 기법인 SUV-MFCC는 한국어에서 유아의 음성인식률이 평균 92.99%으로 높은 인식 결과를 보였다.

SUV-MFCC의 VTLN 와핑 요소 0.8은 높은 피치를 갖는 유아의 음성 특징을 보완하여 와핑을 적용하지 않는 경우엔 약 3.39%, 높은 피치를 약화시키는 경우엔 6.37%의 인식률 향상을 보였다.

또한, 이런 유아의 높은 피치 세분화가 동일 음절을 포함한 단어들을 구분하는데도 많은 도움이 됨을 확인했다.

SUV-MFCC의 대역폭 필터 비교 실험에서 균일 필터가 Constant-Q 필터보다 평균 3.07%의 인식률 향상을 보였다.

특히, 균일 필터가 높은 주파수 성분을 가진 과열음을 포함한 단어의 인식률 향상에 도움이 됨을 확인했다.

실험에서 사용하는 음성 데이터에 대해, SUV-MFCC의 평활화 필터는 삼각필터에 비해 3.49%의 인식률 향상을 보였다.

평활화 필터의 사용이 잡음이 발생하는 다양한 환경에서 잡음에 강건한 알고리즘으로써의 효과를 얻을 수 있음을 기대한다.

본 논문에서 제안한 SUV-MFCC 기법은 유아의 생물학적 특징인 짧은 성대와 높은 피치를 보완하기 위해 MFCC를 변형하는 방법을 제안하였지만, 한국어의 음성학적 특징을 보완하는 방법에 대한 연구도 필요하다. 또한 MFCC를 변형하는 것 뿐만 아니라, 다른 음성 특징인 푸리에 변환, DCT, DWT 등을 한국어 유아 음성 인식 향상을 위해 변형할 수 있으며, 음성 인식기인 HMM의 변형도 가능하다. 추후 실시간 환경에서 존재할 수 있는 다양한 잡음에 강건한 음성 인식기를 개발하여, 다양한 콘텐츠에서 음성 인터페이스를 실질적으로 사용할 수 있는 연구에 힘써야 할 것이다.

참 고 문 헌

[1] 김종훈, 송창우, 김주현, 정경용, 임기욱, 이정현, "음성인식을 이용한 상황정보 기반의 스마트 홈

개인화 서비스", 한국콘텐츠학회논문지, Vol.9, No.11, pp.80-89, 2009.

[2] 권순일, "애니메이션 저작도구를 위한 음성 기반 음향 스케치", 한국콘텐츠학회논문지, Vol.10, No.4, pp.1-9, 2010.

[3] 유재권, 이경미, "한국어에서 성인과 유아의 음성 인식 비교", 한국콘텐츠학회논문지, Vol.11, No.5, pp.138-147, 2011.

[4] A. Potamianos and S. Narayanan, "A review of the acoustic and linguistic properties of children's speech," in proc. of IEEE Multimedia Signal Processing Workshop, 2007.

[5] D. Elenius and M. Blomberg, "Comparing speech recognition for adults and children," in proc. of FONETIK, pp.156-159, 2004.

[6] D. Giuliani and M. Gerosa, "Investigating recognition of children's speech," in proc. of ICASSP, Vol.II, pp.137-140, 2003.

[7] S. Das, D. Nix, and M. Picheny, "Improvements in children's speech recognition performance," in proc. of ICASSP, Vol.I, pp.433-436, 1998.

[8] F. Zheng, G. Zhang, and Z. Song, "Comparison of different implementations of MFCC," Journal of computer science and technology, Vol.16, No.6, pp.582-589, 2001.

[9] S. Umesh and R. Sinha, "A study of filter bank smoothing in MFCC features for recognition of children's speech," IEEE transactions on acoustic. speech and signal processing, Vol.15, No.8, pp.2418-2430, 2007.

[10] 유재권, 이경옥, 이경미, "한국어에서 만 3-5세 유아의 음성 데이터베이스 구축", 한국콘텐츠학회논문지, Vol.12, No.4, pp.52-59, 2012.

[11] S. Young, G. Evermann, M. Gales, T. Hain, D. Kershaw, X. Liu, G. Moore, J. Odell, D. Ollason, D. Povey, V. Valtchev, and P. Woodland, "The HTK Book," Microsoft Corporation and Cambridge University Engineering Department, 2009.

저 자 소 개

유 재 권(Jae-Kwon Yoo)

준회원



- 2009년 2월 : 덕성여자대학교 인터넷정보공학과(공학사)
- 2012년 2월 : 덕성여자대학교 전산정보통신학과(공학석사)
- 2012년 3월 ~ 현재 : (주)파수닷컴 연구원

<관심분야> : 음성인식, 신호처리, 문서보안

이 경 미(Kyoung-Mi Lee)

정회원



- 1993년 2월 : 덕성여자대학교 전산학과(이학사)
- 1996년 2월 : 연세대학교 전산과 학과(이학석사)
- 2001년 12월 : 아이오와주립대학교 전산학과(이학박사)

- 2003년 3월 ~ 현재 : 덕성여자대학교 컴퓨터학과 교수

<관심분야> : 영상처리, 패턴인식, 멀티미디어, HCI