

이종 음성 DB 환경에 강인한 감성 분류 체계에 대한 연구

A Study on Robust Emotion Classification Structure Between Heterogeneous Speech Databases

윤 원 중*, 박 규 식*
(Won-Jung Yoon*, Kyu-Sik Park*)

*단국대학교 컴퓨터학부

(접수일자: 2009년 3월 18일; 수정일자: 2009년 4월 16일; 채택일자: 2009년 5월 18일)

고객센터 (call-center)와 같은 기업환경의 감성인식 시스템은 감성 훈련용 음성과 불특정 고객들의 질의 음성간의 녹취 환경차이로 인해 상당한 시스템 성능 저하와 불안정성을 겪게 된다. 본 논문에서는 이러한 문제점을 극복하기 위해 기존의 전통적인 평상/화남 감성 분류체계를 남녀 성별에 따른 감성별 특성 변화를 적용하여 2단계 분류체계로 확장하였다. 실험 결과, 제안한 방법은 녹취 환경 차이로 인한 시스템 불안정성을 해소할 수 있을 뿐 아니라 약 25% 가까운 인식 성능 개선을 가져올 수 있었다.

핵심용어: 고객센터, 감성분류, 이종 음성 DB

투고분야: 음성처리 분야 (2.5)

The emotion recognition system in commercial environments such as call-center undergoes severe system performance degradation and instability due to the speech characteristic differences between the system training database and the input speech of unspecified customers. In order to alleviate these problems, this paper extends traditional method of emotion recognition of neutral/anger into two-step hierarchical structure by using emotional characteristic changes and differences of male and female. The experimental results indicate that the proposed method provides very stable and successful emotional classification performance about 25% over the traditional method of emotion recognition.

Keywords: Call-center, Emotion classification, Heterogeneous speech DB

ASK subject classification: Speech Signal Processing (2.5)

I. 서론

감성지능 (Emotional Intelligence) 컴퓨팅은 컴퓨터가 학습과 적응을 통하여 인간의 감성을 처리할 수 있는 감성인지 능력을 갖는 것으로 보다 효율적인 인간-컴퓨터 상호작용 (HCI: Human Computer Interaction)을 가능하게 한다. 인간의 감성 정보를 얻는 방법은 얼굴표정, 음성, 몸동작, 심장 박동 수, 체온, 혈압, 뇌파 등 다양한 수단을 통하여 얻을 수 있으나 이 중에서도 음성을 이용한 감성인식 시스템은 마이크로폰을 통한 음성 신호의 입력, 처리 등이 타 매체보다 상대적으로 편리하다는 장점으로 최근에 활발한 연구가 이루어지고 있다. 이러한 음성 감성인식 시스템은 향후 유비쿼터스 (Ubiquitous)

환경에서 음성 정보를 이용하여 상대방의 감성 상태를 알고자 하는 고객센터 (Call Center), 결혼정보 회사, 모바일 콘텐츠 산업 등에서 다양한 형태로 서비스될 수 있을 것이다.

실례로 많은 기업들이 운용하고 있는 고객센터에서는 고객과 상담원들의 대화를 녹취, 관리하면서 상담원들이 고객들에 대해 올바른 응대를 하고 있는지 판단할 수 있는 자료로 활용하고 있다. 예를 들어 불만이 있어 전화한 고객에게는 신입 상담원보다는 경력 많은 유능한 상담원이나 매니저를 연결시켜 상담에 응해준다면 고객의 불만 해결에 도움을 줄 수 있을 것이다. 이러한 기업환경에서 감성인식 시스템은 고객들의 불만을 보다 효율적으로 관리할 수 있는 유용한 시스템이 될 수 있을 것이다.

기존의 음성 감성인식에 대한 연구들로, Dallaert [1]는 피치 윤곽 (pitch contour) 변화를 추출하여 음성 감성상태를 기쁨 (happy), 슬픔 (sad), 화남 (anger), 공포 (fear)

책임저자: 박 규 식 (kspark@dankook.ac.kr)
448-701 경기도 용인시 수지구 죽전동 126번지
단국대학교 정보컴퓨터과학과
(전화: 031-8005-3230; 팩스: 031-8005-3228)

등의 4가지로 분류하여 약 79.5%의 인식률을 달성하였고, Moriyama [2]는 음성 신호의 피치 윤곽과 파워 포락선 (power envelop)을 특징벡터로 사용하여 놀람 (surprise), 화남, 기쁨, 공포, 슬픔 등의 5가지 음성 감성상태를 분류하였으며 이중에서 놀람, 화남, 슬픔 등의 3가지 감성에서 비교적 높은 인식률을 달성할 수 있음을 보였다. A. Nogueiras는 논문 [3]에서 HMM을 이용한 화자종속 방식의 감성인식 시스템으로 놀람, 기쁨, 화남, 공포, 혐오, 슬픔, 평상 등 7개 음성 감성상태를 분류하는데 평균 80%의 인식률을 보이고 있다. 반면 G. Zhou는 논문 [4]에서 TEO (Teager Energy Operator)라는 새로운 특징벡터를 제안하여 평상 음성과 스트레스 음성을 구분하는 흥미로운 연구 결과를 보고하고 있다. 이외에도 미국의 Micro-soft, HP, 일본의 SONY 등의 산업계에서 음성 감성인식 기술을 HCI용 SW나 로봇 등의 응용분야에 적용하기 위한 활발한 연구를 진행하고 있다.

그러나 기존 대부분의 연구가 화자종속 시스템에 대한 연구들이며, 화자독립 시스템에 대한 연구도 동일 환경에서 녹취된 감성음성 DB를 이용하고 있기 때문에, 고객센터와 같이 불특정 다수 화자들의 감성을 인식해야하는 시스템에의 적용이 쉽지 않은 실정이다.

이에 본 논문에서는 불특정 고객들을 대상으로 한 고객센터에서 고객들의 불만관리 시스템에 적용될 수 있는 강인한 감성인식 시스템을 제안하였다. 제안된 시스템은 기존의 전통적인 평상/화남의 이진 감성 분류법 대신 남·녀 성별에 따른 감성별 특성 변화를 이용한 새로운 2단계 감성 분류 방법에 기반하고 있으며, 시스템의 강인성을 평가하기 위해 녹취 환경이 전혀 다른 3개의 감성음성 DB를 사용하였다.

본 논문의 구성은 다음과 같다. 먼저 2장에서는 제안된 음성 감성인식 시스템의 전반적인 구조에 대하여 설명하고, 3장에서는 음성 전처리방법과 시스템에 사용된 특징 벡터들에 대해 간단히 설명한다. 4장에서는 다양한 실험을 통해 제안 시스템의 성능을 비교 분석하고, 마지막으로 5장에서는 결론으로 글을 맺는다.

II. 계층적 감성 분류 체계

실제 고객센터와 같은 기업환경에 적용되는 감성인식 시스템은 감성 훈련을 위한 음성과 불특정 고객들의 질의 음성이 전혀 다른 환경에서 녹취되는 경우가 대부분이다. 이로 인해 기존의 전통적인 평상/화남의 이진

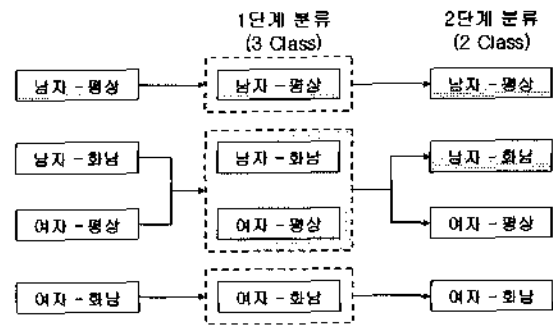


그림 1. 평상/화남 감성의 2단계 분류 체계
Fig. 1. Two-step hierarchical classification for neutral/anger emotion.

감성 분류 방법은 많은 성능 저하와 함께 시스템 불안정성을 초래하게 된다. 이러한 문제점을 극복하기 위해 본 연구에서는 그림 1과 같이 평상/화남의 2가지 감성을 분류할 때, 고객의 남·녀 성별에 따른 감성별 특성 변화를 적용하여 기존의 감성 분류 체계를 2단계 분류 체계로 확장하였다.

먼저 1단계 분류에서는 유사한 특성을 나타내는 남성 화남과 여성 평상을 한 개의 그룹으로 구성하여 총 3개의 class로 분류한 후, 2단계에서는 남·녀 성별을 구분하는 방식으로 최종 평상/화남의 감성인식 결과를 도출한다. 보다 세부적인 분석과 실험 결과는 4.3절에 기술하였다.

III. 감성음성 특징벡터 추출

신뢰성 있는 감성 특징벡터를 추출하기 위한 음성 전처리과정은 그림 2와 같이 프레임 단위의 음성 신호 분할, 해밍 윈도우 (Hamming window) 그리고 끝점 검출로 구성된다.

입력되는 음성 신호는 먼저 20 ms 프레임 단위로 분할되고 이웃한 프레임과 50% 중복되는 Hamming 윈도우를 적용받게 된다. 이후 끝점 검출은 음성 신호에서 음성 구간 (voiced period)과 비-음성 구간을 구별하여 음성 구간에서만 특징벡터를 추출하기 위한 것으로 이는 비-음성 구간에서의 잘못된 음성 분석과 특징벡터 추출로 인해 야기되는 시스템 성능저하를 예방하기 위한 것이다. 본 연구에서 사용된 끝점 검출기는 TEO와 에너지 엔트로피



그림 2. 음성 전처리 과정
Fig. 2. Speech pre-processing.

(entropy)를 이용해 낮은 SNR 환경에서도 비교적 높은 성능을 낼 수 있는 L. Gu의 알고리즘 [5]을 이용하였다.

음성 특징벡터는 매 프레임 단위로 운율적 특징을 갖는 피치, 에너지와 음소 특징을 갖는 MFCC 12차 계수 그리고 각 특징계수의 델타 (Delta) 값을 추출하였고 최종적으로는 각 특징계수들의 평균 (mean)과 표준 편차 (standard deviation)를 구하여 총 56차의 특징벡터를 구성하였다.

IV. 실험 환경 및 결과

4.1. 실험 환경

본 논문에서는 논문 [6]에서 기 구축된 DB와 TV 드라마에서 추출한 DB 그리고 본 연구팀이 직접 구축한 DB의 3가지 감성음성 DB를 이용하여 연구를 진행하였다. 다음 표 1은 각 음성 DB에 대한 간략한 설명이다.

본 논문에서는 표 1과 같이 서로 다른 음성 녹취 환경에서 구축된 3가지의 DB를 이용하여 다음과 같은 주요점들을 두고 실험을 진행하였다.

첫째, DB1을 이용하여 화자독립 환경에서 문장의 독립과 종속이 감성인식 시스템에 미치는 영향에 대한 실험을 진행하였다. 이는 대부분의 기존 연구에 사용된 감성음성 DB들이 모든 감성에 대하여 동일한 문장을 발생하여 구축하는데 반해, 실생활에 적용될 시스템에서는 DB를 구축할 때 사용된 문장이 입력되는 경우가 거의 없기 때문에 문장의 독립/종속 여부가 시스템 성능에 미치는 영향을 파악하기 위함이다.

둘째, 화자독립-문장독립 환경에서 시스템 평가에 사용되는 감성 훈련용 음성과 질의 음성이 동일한 환경에서

녹취된 단일 DB인 실험과 서로 다른 녹취 환경에서 구축된 이중 DB간의 실험을 통하여 기존 연구들의 문제점을 파악하고 화자독립-문장독립 환경에서 강인한 성능을 나타내는 새로운 감성분류 체계를 제안한다.

본 논문에서는 실험을 위해 SVMs 분류기를 사용하였으며, 실험결과는 표 2~5의 1열에 있는 감성을 갖는 질의 음성들에 대한 인식결과를 2~5열에 표시하였다. 또한, 제안된 방법과의 비교를 위한 실험에서는 특징추출 단계에서 추출한 모든 특징들 (pitch, MFCC 그리고 에너지)을 사용하여 단번에 분류하는 전통적인 방법을 사용하였다.

4.2. 실험 I - 문장 독립, 종속이 시스템 성능에 미치는 영향 비교

다음 표 2는 DB1에 대한 실험으로, DB 구축시 문장의 독립/종속 여부가 시스템 성능에 미치는 영향을 파악하기 위한 것이다.

표 2의 화자독립-문장독립 실험은 DB1의 음성 파일 중 평상/화남 감성별 문장이 중복되지 않도록 조절하였기 때문에 실험에 사용되는 음성 파일의 수가 충분하지 않음을 볼 수 있다. 따라서 이들 실험 간의 단순 비교가 어렵기는 하지만, 표 2에서 보듯이 시스템의 성능은 문장의 종속이나 독립의 유형에 따라 15% 가까운 상당한 격차를 가짐을 알 수 있다. 즉, 문장의 독립 혹은 종속의 여부에 따라 시스템 성능에 상당한 영향을 미침을 알 수 있다. 아울러 고객센터 등 실제 기업환경에서 적용되는 감성인식 시스템의 경우, 실용적인 시스템 구축을 위해서는 화자독립-문장독립 시스템을 구축하는 것이 바람직하다고 하겠다.

표 2. DB1에 대한 문장종속과 문장독립 비교 실험결과
Table 2. Comparison result of text-dependent vs text-independent experiment with DB1 test set.

	화자독립 - 문장종속 실험		화자독립 - 문장독립 실험	
	평상	화남	평상	화남
평상	453	47	평상	129
화남	30	470	화남	44
		92.3 %		78.3 %

표 1. 실험에 이용된 DB에 대한 간략한 설명
Table 1. Brief description of DB used in the experiment.

DB 구분	DB1	DB2	DB3
DB에 대한 설명	논문 [6] DB	TV 드라마 DB	본 연구팀 DB
녹음 인원	이미추어 연구단원 남·녀 각 15명	전문 연기자 남·녀 각 30명	이미추어 연구단원 남·녀 각 20명
발성 문장	모든 감성에 동일한 45 문장	각기 다른 2,000 문장	각 감성별 100 문장
음성 파일 수	2,700 개	2,000 개	8,000 개
샘플링 주파수	8 kHz	8 kHz	8 kHz
특징	화자 독립 문장 종속	화자 독립 문장 독립	화자 독립 문장 독립

4.3. 실험 II - 이중 훈련용 음성 DB와 질의 음성의 영향 비교

실제 기업환경에 적용될 수 있는 화자독립-문장독립 감성인식 시스템의 성능평가를 위해 훈련용 음성과 질의 음성이 동일한 녹취 환경인 경우와 서로 다른 경우에 대

한 실험을 진행하였다. 이를 위해 TV 드라마 대사를 녹음한 DB2와 본 연구팀이 준비한 DB3를 사용하였다.

TV 드라마 대사를 녹음한 DB2는 남·녀 각 30명의 연기자가 TV 드라마에서 발성한 대사를 녹음한 것으로 총 2,000개의 음성 파일로 구성되며, 각 화자별로 중복되는 문장이 없다는 것이 특징이다.

본 연구팀의 DB3는 TV 드라마 대사들을 참고하여 비교적 감성이 잘 표현될 수 있는 2~3초 길이의 문장으로 평상/화남 각 감성별 100문장으로 구성하였으며, 아마추어 연극단원 남·녀 각 20명이 발성한 총 8,000개의 음성 파일로 구성된다. 또한, 실험결과를 쉽게 비교할 수 있도록 이후의 실험에서 질의 음성 파일의 수는 1,000개씩으로 제한하였다.

다음의 표 3은 훈련용 음성과 질의 음성이 동일한 녹취 환경인 감성 음성 DB2와 DB3에 대한 실험 결과이다.

표 3. DB2와 DB3에 대한 화자독립-문장독립 실험결과
Table 3. Experimental result of speaker-independent, text-independent case with DB2 and DB3 test set.

	DB2		DB3		
	평상	화남		평상	화남
평상	433	67	평상	493	7
화남	112	388	화남	20	480
		82.1 %			97.3 %

표 3에서 보듯이 DB2의 경우 표 2의 DB1 결과와 큰 차이가 없음을 알 수 있으나, 동일한 특징과 알고리즘을 적용했을 경우에도 DB3가 15%이상의 높은 성공률 향상을 가져왔다. 이는 DB3를 구축할 때 비교적 감성이 잘 표현될 수 있는 문장과 더불어 아마추어 연기자들의 감성 표현이 잘 되어 DB3가 안정적으로 구축된 결과라고 분석할 수 있다. 이는 감성음성 DB 구축 방법 또한 시스템의 성능에 상당한 영향을 미칠 수 있음을 의미한다.

다음의 표 4는 화자독립-문장독립 환경에서 감성 훈련용 음성과 질의 음성의 녹취 환경이 서로 다른 경우에 대한 실험 결과이다. 이는 실제 기업환경에 적용되는 감성인식 시스템의 경우 감성 훈련용 음성과 고객의 질의 음성이 서로 다른 녹취 환경을 갖기 때문에 반드시 고려되어야 할 사항이다. 표 4의 실험 역시 표 2, 3의 실험과 동일하게 특징추출 단계에서 추출한 모든 특징들을 사용하여 단번에 분류하는 전통적인 방법을 사용하였다. 표 4에서 괄호 안의 (D)는 감성 훈련용 음성의 종류를 말하며 (Q)는 시스템 평가에 사용된 질의 음성의 종류를 말한다.

표 4. 훈련용 음성과 질의 음성의 녹취 환경이 다른 경우의 실험 결과

Table 4. Experimental result when training speech and query speech are recorded in different environment.

DB1 (D) vs DB2 (Q)			DB1 (D) vs DB3 (Q)		
	평상	화남		평상	화남
평상	368	132	평상	343	157
화남	87	413	화남	7	493
		78.1 %			83.6 %

DB2 (D) vs DB1 (Q)			DB2 (D) vs DB3 (Q)		
	평상	화남		평상	화남
평상	424	76	평상	420	80
화남	217	283	화남	70	430
		70.7 %			85.0 %

DB3 (D) vs DB1 (Q)			DB3 (D) vs DB2 (Q)		
	평상	화남		평상	화남
평상	500	0	평상	490	10
화남	490	10	화남	385	115
		51.0 %			60.5 %

표 4에서 보듯이 실험 결과는 50%~85%까지 다양한 성능을 나타내었다. 우선, DB1의 평상 음성에 대한 훈련 결과는 DB2, DB3의 평상 음성과 74%, 68%의 유사함을 보였고, DB1의 화남 음성에 대한 훈련 결과는 DB2, DB3의 화남 음성과 82%, 99%의 유사함을 보였다. 다음으로 DB2의 평상 음성에 대한 훈련 결과는 DB1, DB3의 평상 음성과 84%의 유사함을 보였고, DB2의 화남 음성에 대한 훈련 결과는 DB1, DB3의 화남 음성과 57%, 86%의 유사함을 보였다. 안정적으로 잘 구축된 DB3의 음성은 서로 특성이 다른 DB1과 DB2의 훈련결과 모두와 높은 인식률을 나타내었다.

그러나 이러한 DB3를 훈련용 음성으로 사용했을 때는 한쪽 감성, 즉, 평상 감성으로의 쓸림현상이 나타나며 전혀 예측할 수 없는 현상을 나타낸다. 이는 아무리 잘 구축된 DB를 훈련용 음성으로 사용한다고 해도 어떤 환경에서 녹취된 (어떤 특성을 가진) 음성이 들어오는데 따라 시스템의 성능이 불안정해질 수 있다는 것을 의미한다. 이러한 감성 쓸림 현상은 실용적인 감성인식 시스템의 구축에 가장 큰 성능 저하 요소로 작용할 수밖에 없다.

다음의 그림 3은 표 4의 편향된 감성 쓸림 현상을 설명하기 위해 감성인식에서 가장 중요한 특징벡터로 알려져 있는 pitch mean의 남·녀 성별에 따른 감성별 특성 변화를 나타낸 것으로, 남·녀 각 10명씩의 화자가 평상/화남에 대하여 30문장씩을 발성한 결과의 pitch mean 값을 보여주고 있다. 30의 배수마다 눈금선을 표시하여 각 화

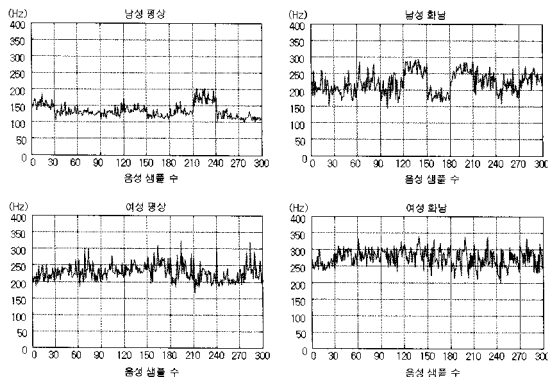


그림 3. 남·녀 성별에 따른 Pitch mean의 감성별 특성 변화 비교
Fig. 3. Comparison of emotional characteristic changes in pitch mean between male and female.

자별 pitch mean값을 비교할 수 있도록 하였으며, 세로 축은 pitch mean frequency를 나타낸다.

사용된 감성음성 DB에 따라 약간씩의 차이는 있지만, 그림 3에서 보듯이 pitch mean 값의 범위는 남성 평상에 대해 가장 낮고(약 160 Hz이하), 남성 화남과 여성 평상의 범위가 유사한 영역에 존재(약 160 Hz이상, 250 Hz이하)하며, 여성 화남이 가장 높은 범위에 존재(약 250 Hz 이상)한다. 만약 훈련용 음성과 질의 음성의 pitch mean 값의 범위가 유사하다면 큰 문제가 없겠지만, 녹취 환경 및 화자의 차이로 인해 값의 범위에 현격한 차이가 있다면 이는 표 4의 하단에 보이는 2개의 실험결과처럼 감성 인식 결과가 한쪽으로 치우치게 나타날 것이다. 또한, 질의 음성의 남·녀 성별을 고려하지 않고 평상/화남의 감성 분류를 할 경우, 남성 화남과 여성 평상의 특징이 상당히 유사한 값의 영역에 존재하고 있기 때문에 상당한 오 분류를 야기하기도 한다.

이러한 문제를 해결하기 위해 본 연구에서는 그림 1과 같이 남·녀 성별에 따른 감성별 특성 변화를 감안한 2단계 평상/화남 감성 분류법을 사용하였다.

먼저 1단계 분류에서는 pitch mean을 기준으로 질의 음성을 (남성 화남-여성 평상), (남성 평상), (여성 화남) 등 3개의 class로 분류하고, 2단계에서는 MFCC와 에너지를 이용해 남·녀 성별을 구분해 남성 화남과 여성 평상 음성을 분류한다. 표 5는 본 연구에서 제안한 2단계 감성 인식 방법으로 실험한 경우의 평상/화남 감성 분류 결과를 나타낸다.

표 4와 비교해 볼 때 인식 감성이 한쪽으로 쏠리는 현상이 없음을 볼 수 있을 뿐만 아니라, 실험 결과 역시 51% → 86.3%, 60.5% → 75.1%로 평균 25%의 성능 개선이 있음을 알 수 있다. 즉, 본 연구의 남·녀 성별 특징 변화를 감안한 2단계 감성 분류를 통해 감성 훈련용 음성과 질의 음성의

표 5. 제안된 방법에 의한 감성 분류 결과

Table 5. Emotion classification results by the proposed method.

DB3 (D) vs DB1 (Q)				
	남성 평상	남성 화남	여성 평상	여성 화남
남성 평상	195	55		0
남성 화남	0	205	29	45
여성 평상		1	220	
여성 화남	0	7		243
				86.3 %

DB3 (D) vs DB2 (Q)				
	남성 평상	남성 화남	여성 평상	여성 화남
남성 평상	198	52		0
남성 화남	30	165	29	63
여성 평상		25	188	
여성 화남	0	50		200
				75.1 %

녹취 환경이 서로 다른 경우에 발생하던 편향된 인식 결과 문제를 해결할 수 있으며, 이와 더불어 상당한 인식 성능 개선 또한 가져올 수 있음을 알 수 있다.

V. 결론

본 논문은 고객센터와 같은 기업환경에서 고객들의 불만관리 시스템에 적용할 수 있는 평상/화남의 2가지 감성에 대한 강한 감성분류 시스템을 제안하였다. 실제 고객센터에서 사용되는 감성인식 시스템은 감성 훈련용 음성과 불특정 고객들의 질의 음성이 전혀 다른 환경에서 녹취되는 경우가 대부분이기 때문에 기존의 전통적인 평상/화남의 이진 감성 분류 방법은 많은 성능 저하와 함께 시스템 불안정성을 초래하게 된다. 이러한 문제점을 극복하기 위해 본 연구에서는 평상/화남 2가지 감성을 분류할 때 남·녀 성별에 따른 감성별 특성 변화를 적용하여 기존의 감성 분류 체계를 2단계 분류 체계로 확장하였다. 실험 결과 제안한 방법은 녹취 환경 차이로 인한 편향된 인식 결과의 해결과 더불어 상당한 인식 성능 개선 또한 가져올 수 있음을 알 수 있었다.

제안 시스템은 특히 고객센터, 소방서, 경찰서 등과 같이 불특정 다수로부터 불만이나 공포의 감성을 인식하고 빠르게 대응해야하는 곳에서 유용하게 활용될 수 있을 것으로 예상된다. 향후에는 2가지 감성에 대한 연구를 확장하여 4가지 감성에 대하여 이중 DB간에도 강인하게 동작하는 시스템에 대하여 연구하고자 한다.

감사의 글

본 논문은 2008학년도 단국대학교 대학연구비의 지원으로 연구되었음.

참고 문헌

1. F. Dellaert, T. Polzin, and A. Waibel, "Recognizing Emotion in Speech," *In Proc. International Conf. on Spoken Language Processing*, pp. 1970-1973, Oct., 1996.
2. T. Moriyama and Oozwa, "Emotion Recognition and Synthesis System on Speech," *IEEE International Conference on Multi-media Computing and Systems*, vol. 1, pp. 840-844, July, 1999.
3. A. Nogueiras, A. Moreno, A. Bonafonte, and J. B. Marino, "Speech Emotion Recognition Using Hidden Markov Models," *In Eurospeech 2001 Poster Proceedings*, pp. 2679-2682, Sep., 2001.
4. Guojun Zhou, John H. L. Hansen, and James F. Kaiser, "Nonlinear Feature Based Classification of Speech Under Stress," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 3, pp. 201-216, 2001.
5. Lingyun Gu and Stephen A. Zahorian, "A New Robust Algorithm for Isolated Word Endpoint Detection," *International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 4161, May, 2002.
6. 강봉석, "음성 신호를 이용한 문장독립 감정 인식 시스템," 석사 학위 논문, 연세대학교, 2001.

저자 약력

•윤 원 중 (Won-Jung Yoon)

한국음향학회지 제23권 7호 참조

•박 규 식 (Kyu-Sik Park)

한국음향학회지 제23권 7호 참조