

소리의 물리적 속성을 활용한 소리에 대한 감성인식 모델 개발 Development of Emotion-Recognition Model for Sound's Psycho-acoustic Parameters

*최영임¹, #정성수², 이성준, 최인묵, 박연규, 강대임²

*Y. I. Choi¹, S. S. Jung(jjs@kriss.re.kr)¹, S. Lee¹, I. M. Choi¹, Y. K. Park¹, D. I. Kang²

¹한국표준과학연구원 기반표준본부, ²한국표준과학연구원 휴먼인지환경사업본부

Key words : psycho-acoustic parameter, emotion recognition

1. 서론

많은 연구들은 유비쿼터스 환경에서 인간과 기계가 상호작용하는 방식은 기존의 컴퓨터나 기계 중심 설계에서 사용자가 중심이 되는 인간중심의 설계로 변화 되어야 한다는 것을 강조한다. 이러한 인간 중심 설계로의 전환에서의 핵심은 사용자의 정서상태를 민감하게 인식하는 것에 있다. 즉, 기계는 사용자의 요구 단순한 반응을 보이는 것보다 오히려 사용자의 행동 특히, 인간의 정서적 행동에서의 미묘함과 정서의 변화를 탐지할 수 있어야 하고, 그러한 정보에 기초된 상호작용을 할 수 있는 능력을 지녀야 한다는 것을 의미한다[1].

최근에는 인간의 감성을 모방하는 로봇을 개발하거나, 감성과 관련된 특성들을 추출하여 감성을 인식하는 연구들이 수행되고 있다. 이러한 연구들은 다양한 감정을 나타내는 인간의 얼굴표정의 특징 즉, 눈동자의 움직임이나 눈썹 및 입의 모양 등을 인식하여 감성을 인식하고자 한다. 또한 다양한 감정 표출 시 달라지는 맥박이나 심전도 및 열감과 같은 생리적 신호들을 추출하여 인간의 감성을 인식하는 연구도 수행되고 있을 뿐만 아니라 감성을 수반하는 인간의 목소리나 음악으로부터의 감성인식 연구도 수행되고 있다. 그러나 인간의 얼굴인식이나 생체신호 및 음성 정보를 통한 감성인식 연구에 비하여 생활 속에서 듣게 되는 자연음에 대한 감성인식에 대한 연구는 거의 수행되지 않은 것이 현실이다. 그러나 인간은 매 순간에 생활 소리들에 노출되며, 그러한 소리를 통하여 다양한 감성을 경험한다. 자동차 경적소리나 알람소리 및 유리컵이 깨지는 등의 소리는 놀람이나 짜증의 감성을 유발하는 반면 새가 지저귀는 소리나 시냇물이 흐르는 소리 등은 편안함고 이완

된 감성을 유발할 수 있다.

또한 인간의 감각기관은 각 감각기관으로부터 들어오는 정보를 통합하여 사물을 지각하는 뛰어난 능력인 다중감각정보 활용능력을 지니고 있다. 예를 들면, 배경지식이 부족한 아동들에게 동화를 들려주는 상황에서 동화의 내용과 연관된 그림을 함께 제시해 주었을 때 아동의 작업기억능력이 향상될 수 있다. 이처럼 인간은 단일감각을 통한 정보해석보다는 두 가지 이상의 감각기관이 동시에 정보를 해석하여 인지하는 다중감각처리를 빈번하게 수행하므로, 따라서 청각적 정보에 대한 감성인식 모델 개발은 시각적 정보에 의존한 기존의 감성인식 모델의 성능향상에 기여할 수 있다.

따라서 이 연구는 일상생활에서 흔하게 들을 수 있는 생활소리 및 자연음으로부터 소리의 심리음향특성을 추출하여, 소리로 부터 유발되는 감성이 무엇인지를 판별해내는 함수식을 산출해내는 것을 목적으로 한다.

2. 방법

· **실험참가자**: D시에 소재한 대학교에 재학 중인 대학생 140명을 실험참가자로 선정하였다.

· **실험재료**: IADS-2(The International Affective Digitized Sounds 2nd Edition)[2]를 정서를 유발하기 위한 실험재료로 사용하였다. IADS-2는 unpleasure에서 부터 pleasure까지 분포된 Valence차원과 calm에서 부터 excited까지 분포된 Arousal차원의 2차원 공간에 분포되어 있으며 6초 동안 제시되도록 통제된 167개의 표준화된 감성유발 소리자극이다. 이 중 두 차원에 고르게 분포되어 있는 44개의 소리를 선정하여 감성유발자극으로 사용하였다.

· **실험절차**: 참가자들은 7명이 한 팀이 되어 실험에 참가하였다. 참가자들은 실험진행자가 불러주

는 소리번호를 받아 적고, 6초 동안 제시되는 자극을 청취하였다. 소리가 끝난 직후, 소리로 부터 유발되어 느껴지는 정서를 기쁨, 슬픔, 분노, 공포, 혐오의 5가지 기본정서 중 해당하는 정서에 평정하도록 하였다.

· 심리음향특성 추출: 각 소리에 대해 Zwicker[3]의 심리음향특성 중 Loudness, Sharpness, Roughness, Fluctuation Strength의 4개 요인을 추출하였다. 소리의 샘플링과 녹음방식에서의 차이를 최소화하기 위하여 샘플링은 44200Hz로 그리고 녹음방식은 mono음으로 통제하였다.

3. 결과

제시된 소리에 대해 평정된 감성의 빈도를 분석한 결과, 기쁨 8개, 슬픔 8개, 분노 10개, 공포 7개, 혐오 11개로 나타났다. 심리음향특성들로 이러한 소리에 대한 감성을 판별할 수 있는 함수식을 산출하기 위하여, 4가지 심리음향 특성을 독립변인으로 하고 5가지 정서를 집단변인으로 하여 판별분석을 실시하였다. 4가지 심리음향특성으로 다섯 가지 정서를 판별할 수 있는 판별분류함수를 산출하였다.

$$Happy = -222.315 + (299.48 * Loudness) + (-34.33 * Sharpness) + (55.156 * Fluctuation Strength) + (-17.98 * Roughness)$$

$$Sad = -229.90 + (310.37 * Loudness) + (-44.46 * Sharpness) + (45.55 * Fluctuation Strength) + (-13.09 * Roughness)$$

$$Anger = -244.40 + (318.22 * Loudness) + (-43.05 * Sharpness) + (53.87 * Fluctuation Strength) + (-14.00 * Roughness)$$

$$Fear = -221.45 + (310.51 * Loudness) + (-59.20 * Sharpness) + (23.24 * Fluctuation Strength) + (1.55 * Roughness)$$

$$Disgust = -196.06 + (287.71 * Loudness) + (-36.24 * Sharpness) + (36.15 * Fluctuation Strength) + (1.11 * Roughness)$$

이 판별분류함수를 이용하여 44개의 소리의 감성을 판별한 결과, 표 1에 제시된 바와 같이, 기쁨은 6개(75%), 슬픔은 4개(50%), 분노는 10개(100%), 공포는 7개(100%), 혐오는 11개(100%)의 자극이 실험참가자들이 평정한 정서와 일치된 것으로 판정되었으며, 전체 86.4%의 판별정확율을 나타내었다. 이렇게 판별된 함수의 타당성을 확보하기 위하여 교차타당화 분석을 실시한 결과 기쁨은 5개, 슬픔은 3개, 분노는 9개, 공포는 7개, 혐오는 9개가 정확한 정서로 판별되어 전체 75.8%의 판별정확율을 보였다. 소리자극에 대한 판별분포를 그림 1에 제시하였다.

4. 논의

많은 연구가 수행되지 않은 일상 생활소리음을 감성유발자극으로 선정하여 4가지 심리음향 특성만을 추출하여 5가지 정서를 75% 이상 정확하게 판별할 수 있는 함수를 개발하였다는 점에서 연구의 의의를 찾을 수 있다. 이러한 연구결과는 얼굴 및 음성 등에 대한 감성인식 연구의 정확도를 향상에 기여할 수 있을 뿐만 아니라 청각정보를 사용하는 다양한 기계의 감성기술 도입에 기여할 수 있을 것으로 기대된다.

Table 1. Classification Results

Emotion	Predicted Group Membership					Total
	Happy	Sad	Anger	Fear	Disgust	
Happy	6	1	1	0	0	8
Sad	2	4	2	0	0	8
Anger	0	0	10	0	0	10
Fear	0	0	0	7	0	7
Disgust	0	0	0	0	11	11
Happy	75.0	12.5	12.5	.0	.0	100.0
Sad	25.0	50.0	25.0	.0	.0	100.0
Anger	.0	.0	100.0	.0	.0	100.0
Fear	.0	.0	.0	100.0	.0	100.0
Disgust	.0	.0	.0	.0	100.0	100.0

* 86.4% of original grouped cases correctly classified.

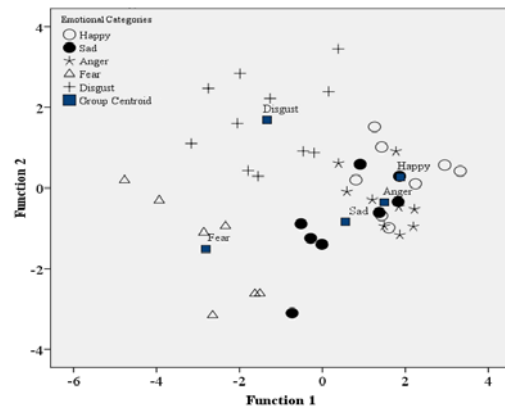


Figure 1. Canonical Discriminant Functions

후기

"이 논문은 2010년도 정부(교육과학기술부)의 재원으로 휴먼인지환경사업본부-신기술융합형 성장동력사업의 지원을 받아 수행된 연구임(No. 2010K001125)."

참고문헌

- Zhihong, Z., Maja, P., Glenn I. R. & Thomas, S. H., (2007). A survey of affect recognition methods: Audio, visual and spontaneous expressions. ICMI'07, November 12-15, Nagoya, Aichi, Japan.