



음성인식 분석에 따른 감정 패턴에 관한 연구

A Study on the Emotion Patterns based on Speech Recognition Analysis

저자 (Authors)	서은주, 이기영 Eun-Joo Seo, Ki-Young Lee
출처 (Source)	대한전자공학회 학술대회 , 2012.6, 718-720 (3 pages)
발행처 (Publisher)	대한전자공학회 THE INSTITUTE OF ELECTRONICS ENGINEERS OF KOREA
URL	http://www.dbpia.co.kr/Article/NODE02275148
APA Style	서은주, 이기영 (2012). 음성인식 분석에 따른 감정 패턴에 관한 연구. 대한전자공학회 학술대회, 718-720.
이용정보 (Accessed)	금오공과대학교 202.31.143.*** 2019/03/07 13:48 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독 계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

음성인식 분석에 따른 감정 패턴에 관한 연구

*서은주, 이기영

을지대학교 의료IT마케팅학과

e-mail : eunjoo-grace@hanmail.net, kylee@eulji.ac.kr

A Study on the Emotion Patterns based on Speech Recognition Analysis

*Eun-Joo Seo, Ki-Young Lee

Department of Medical IT and Marketing

Eulji University

Abstract

Recently, According to the use field of voice recognition technology in the mobile market expands, expectation for the provision of contents to consider sentiment or intent through the voice is growing. In this paper, We proposed that voice signal of human is analyzed to predict the current emotional state. In addition, we design interface that the correlation between emotional states and bio-signals to realize the human emotion more detail.

에 기반 한 특정 단어에 대한)확률값 등의 6가지 정보들로부터 최종 감정상태 추출 알고리즘을 정의하였다. 그러나 이는 우리가 자신의 감정을 잘 표현하는 상황에서만 적용할 수 있다는 한계에 부딪히게 된다. 예를 들어 사용자가 화가 났음에도 불구하고 내성적인 성격 탓에 감정을 잘 표현하지 못하고 있는 상황에서는 감정분류가 쉽지 않다.

이에 본 논문에서는 음성인식뿐만 아니라 자연스롭게 발생하는 감정을 파악하기 위해 생체 신호 분석을 융합시킨 알고리즘을 제시하였다. 이를 통해 조금 더 세분화된 감정분류가 가능할 것이라고 판단된다.

II. 본론

I. 서론

현재 IT시장은 사용자의 직접적인 입력에만 의존한 콘텐츠 제공에서 사용자의 정서 및 감성을 고려한 콘텐츠 제공에 대한 기대가 커지고 있다. 이에 따라 감정 정보를 분석하기 위한 연구는 감정을 토대로 표현되는 음성인식에 관심이 집중되고 있다. 보통 사람들은 대화에서 표정 또는 음성을 통해 상대방의 감정을 파악할 수 있다. 특히 음성은 직접 대면하지 않는 상황 속에서도 상대방의 감정을 쉽게 알 수 있기 때문에 감정분류에 있어서 활용도가 높다.

기존 음성을 통한 감정분류 연구에서는 음성신호를 통해 얻어지는 값들의 평균값과 분산값, (데이터베이스

2.1 음성신호를 통한 감정분류

음성으로부터 인간의 감정 상태를 파악하는 연구는 꾸준히 발전되어왔다. 그 중 후쿠다(Fukuda)의 연구에서는 보통상태와 화가 난 각 상태에서 수집된 음높이, 음성 에너지와 같은 음성 데이터를 이용하여 음성의 감정 상태를 추출하였다. 그러나 이 연구[1]에서 음성 데이터는 외부 환경적 요인이 작용하게 되면 감정 분류에 정확성이 떨어진다는 판단 하에 발화의 내용면을 파악하였다. 그래서 이를 감정 인식의 준거로 활용하도록 함으로써 신뢰성을 높였다. 각각의 음성 데이터로부터 평균값과 표준편차를 구하고, 발화의 내용에 대해서는 미리 구축해 둔 데이터베이스를 통해 특정 단어에 대한 확률 값을 구하였다. 이렇게 얻어진 값들을 보통상태의 값들과 함께 식 (1)을 통해 계산하여 AER(감정분류계수)를 얻었다.

$$\begin{aligned}
 AER = & a_0 (\overline{P - Pr}) / \overline{Pr} + a_1 (\overline{E - Er}) / \overline{Er} \\
 & + a_2 (\partial P - \partial Pr) / \partial Pr + a_4 (\partial E - \partial Er) / \partial Er \\
 & + a_5 (\overline{D - Dr}) / \overline{Dr} + a_6 Pw
 \end{aligned} \quad (1)$$

입력된 음성신호의 음높이 평균값 \overline{P} , 음성에너지 평균값 \overline{E} , 음높이 분산 $\partial^2 P$, 음성에너지 분산을 $\partial^2 E$, 음소길이 평균값을 \overline{D} , 특정 단어에 대한 확률값을 Pw , 대량의 음성데이터로부터 구한 평상시 상태에서의 음높이 평균값 \overline{Pr} , 음성에너지 평균값 \overline{Er} , 음높이 분산 $\partial^2 Pr$, 음성에너지 분산을 $\partial^2 Er$, 음소길이 평균값을 \overline{Dr} 라고 하였을 때의 식이다. $a_0 \sim a_6$ 은 각 파라미터에 대한 가중치이며, 이를 통해 AER을 산출한 후 각 감정에 대해 미리 설정해 두었던 임계치로 감정 상태를 분류하게 된다.

2.2 문장을 통한 감정분류

한국어 대화체 문장으로 감정을 분류하는 연구[2]에서는 이 모티콘과 기호문자에 가장 높은 확률을 부여했기 때문에 일반적인 자연어 처리에서 적합하지 않았다. 또한 어절만을 자질로 사용함에 있어서 어절 당 형태적 변형이 빈번하게 일어나는 한국어를 분석하기 위한 다양한 데이터베이스 구축의 한계를 가진다.

또 다른 논문[3]에서는 감정을 분류하기 위해 발화 문장을 문장자질, 선행적 자질, 문맥자질로 나누어 Support Vector Machine을 적용하였다. 이러한 감정을 분류하기 위해 사용된 분류자질은 본 연구와 유사하다. 그러나 본 연구는 분류자질들과 앞서 음성을 통해 얻은 감정분류계수들의 Naive Bayes 확률을 거쳐 감정분류가 이루어진다는 면에서는 차이점을 가진다.

2.3 생체신호를 통한 감정분류

간편하게 생체신호를 얻어낼 수 있는 맥파와 피부전도도를 이용해서 감정을 분류하는 연구는 활발히 진행 중이다. 맥파를 이용한 논문[4]에서는 맥파를 이용하여 HRV를 검출하였다. 또 얻은 HRV를 Fast Fourier Transform하여 얻어진 그래프를 통해 LF/HF의 비율을 얻어내어 사용자의 부정적 상태와 긍정적 상태를 판별해 낼 수 있는 근거로 사용하였다.

피부 전도도를 사용한 연구[5]에서는 진폭을 변수로 사용하여 지정된 구간 내의 영교차율과 평가 구간의 평균을 구하여 각 심리상태를 알고리즘을 통해 긴장/이완 상태로 구별해내고 있다.

III. 설계 및 구현

사용자마다 음성을 통해 나이와 성별에 대한 정보를 받아

데이터베이스에 저장하도록 하고, 사용자로부터 입력받은 음성 데이터를 분석하여 사용자의 감정을 7가지의 감정상태(보통, 화남, 슬픔, 즐거움, 기쁨, 우울, 당황)로 분류되도록 설계하였다. 또한 입력된 생체 신호 데이터를 결합하여 7가지의 감정 상태를 보다 더 세밀하게 분류되도록 설계하였다.

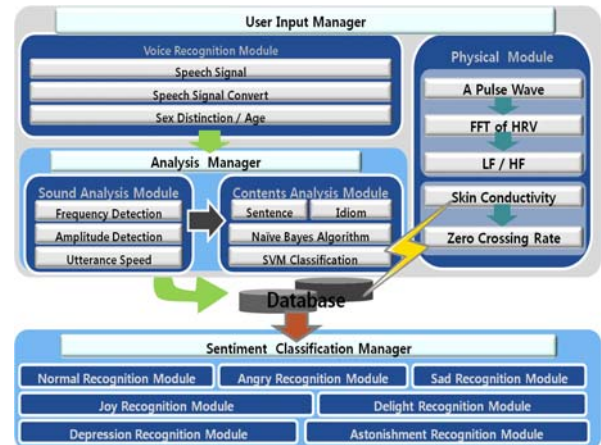


그림 1. 시스템 구조도

전체적인 시스템은 그림 1과 같이 Input Management, Analysis Management, Sentiment Classification Management로 이루어진다.

Input Management는 사용자 개개인으로부터 음성 데이터와, 생체신호 데이터를 입력받는다. 음성 신호는 먼저 전 처리 과정 후 진폭과 주파수에 대해서는 각각의 평균값과 표준편차 값을 얻고, 데이터 값이 존재하는 시간영역에 대해서는 발화속도를 얻는다. 분석하여 얻은 데이터 값들은 감정 분류 계산식을 이용하여 수치로 산정되어지고, 이를 통해 7가지의 감정상태 중 가능성 있는 감정의 단계를 추정하게 된다.

Content Analysis Module에서는 모바일 단말기를 통해 얻은 문장을 데이터베이스에 저장되어 있는 데이터와 비교하여 그림 2와 같은 관용어 문구일 경우 감정 결정이 이루어진다. 그렇지 않을 경우에는 음성단계의 감정분류 임계치와 문장 정보를 결합하여 Naive Bayes Algorithm을 통해 감정 결정이 이루어진다.

Physical Module에서는 맥파 신호로부터 피크점을 검출하고, 피크점들의 간격들의 변화량을 통해 HRV를 추출한다. HRV 그래프를 FFT(Fast Fourier Transform)함으로써 얻은 주파수 영역의 그래프를 통해 HF/LF비율에 따라 긍정적 상태와 부정적 상태로 분류한다. 피부전도도의 경우 구간 내 진폭의 평균값을 통해 각 심리상태의 분석 도구로 활용하였다. 현재 상태에서 받은 피부전도도 값은 평상시 피부전도도의 값과 비교하여 안정 상태인 평온, 집중과 불안정 상태인 긴장, 우울로 분류하였다.

Voice Recognition Module과 Physical Module을 통해 얻은

출력값들은 데이터베이스에 저장되어진 Algorithm을 통해서 최종 감정분류가 이루어진다

IV. 결론

본 논문에서는 모바일 단말기로부터 받은 음성 및 문장 데이터와 블루투스를 통해 받은 생체신호 데이터를 결합하여 사용자의 감정을 추정하는 시스템을 설계하였다. 최근 한 외국계의 기업에서는 음성을 통해 사용자의 지시를 받고 사용자의 의도까지 파악하여 적절한 콘텐츠를 제공하는 제품을 출시하여 각광받고 있다. 그러나 현재 모바일에 적용된 음성인식 기술은 사용자가 주는 언어적인 부분에 의존하고 있기 때문에 향후 비언어적인 음성신호나 생체신호를 통해 사용자의 감정까지 고려한다면 인간과 기계의 피드백을 강화시켜주는 역할을 할 것이다.

참고문헌

- [1] 한우진, 최도연, "음성을 이용한 감정인식 시스템 및 감정인식 방법", 공개특허 10-2004-0038419, 공개일자 2004.
- [2] 문현구, 장병탁, "채팅 텍스트로부터의 화자 감정 상태 학습", 한국정보과학회 학술발표논문집, 제28권, 제1호, pp. 340-342, 2001.
- [3] 강상우, 박홍민, 서정연, "대화 시스템을 위한 사용자 발화 문장의 감정 분류", 한국인지과학회논문지, 제21권, 제4호, pp. 459-480, 2010.
- [4] 이현민, 김동진, "맥파를 이용한 감성평가 및 바이오피드백 시스템 개발", 대한전기학회논문지, 제57권, 제6호, pp. 1087-1094, 2008.
- [5] 전기환, 최형진, "생체정보를 이용한 유비쿼터스 심리상태 인식 모델 연구", 한국통신학회논문지, 제35권, 제2호, pp. 232-243, 2010.