

스마트폰에서 음성과 컨텍스트 기반 감정인식 융합을 활용한 감정인식 시스템 설계

조성진*, 이성호°, 이승룡*

*°경희대학교 컴퓨터공학과

e-mail: {sjcho, sylee}@oslab.khu.ac.kr*, sholee@oslab.khu.ac.kr°

Design of Emotion Recognition system utilizing fusion of Speech and Context based emotion recognition in Smartphone

Seong Jin Cho*, Seongho Lee°, Sungyoung Lee*

*°Dept. of Computer Engineering, Kyung Hee University, Korea

● 요약 ●

최근 스마트폰 환경에서 제공되는 수많은 서비스들은 일률적으로 소비자에게 단방향으로 제공되던 예전과 달리 사용자마다 개인화된 서비스 제공을 통해, 더욱 효율적으로 서비스를 제공하려는 시도들이 이루어지고 있다. 그 중에서 감정인식을 이용한 연구는 사용자에게 최적화된 개인화 서비스 제공을 위해 사용자의 감정을 인식하여 사용자가 느끼는 감정에 맞는 서비스를 제공함으로써 보다 몰입감을 느낄 수 있도록 하여 결과적으로 특정 서비스의 이용을 유도하도록 할 수 있다. 본 논문에서는 사용자 선호도와 컨텍스트 정보를 활용하여 사용자의 감정을 추출하고 이를 음성기반 감정인식과 융합하여 그 정확도를 높이고 실제 서비스에서 활용할 수 있는 시스템 설계를 제안한다. 제안하는 시스템은 사용자 선호도와 컨텍스트 인식으로 감정을 판단했을 경우의 오류를 음성을 통한 감정인식으로 보완하며, 사용자가 감정인식 시스템을 활용하기 위한 비용을 최소화한다. 제안하는 시스템은 스마트폰에서 사용자 감정을 이용한 애플리케이션이나 추천서비스 등에서 활용이 가능하다.

키워드: 감정인식(emotion recognition), 스마트폰(smartphone)

I. 서론

최근 스마트폰에는 다양한 센서들이 탑재되어 센서 네트워크를 활용한 유비쿼터스 서비스가 가능해지고 있다. 그러나 수많은 서비스들이 사용자들에게 일률적으로 제공되고 있으며 개개의 사용자들이 모든 서비스들을 경험하는 것은 불가능하다. 그러므로 사용자들은 과도하게 공급되는 서비스들 중에 자신에게 맞는 개인화된 서비스를 제공받기를 원한다. 이를 위해 스마트폰에 내장된 센서를 활용하여 사용자의 감정을 인식하고, 이를 기반으로 사용자 개개인의 감정을 반영한 서비스 제공이 시도되고 있다. 그러나 기존에는 스마트폰을 활용하여 감정을 인식하는 데 있어서 사용자의 입력에 대한 의존도가 크거나 센서를 사용자가 직접 제어하여 감정을 인식해야 하는 등 실제로 스마트폰에서 활용하기에는 그 실용성이 미흡하였다. 제안하는 시스템은 감정인식을 위해 사용자의 선호도정보와 컨텍스트 정보를 활용하며 추출한 감정의 신뢰도를 높이기 위해 음성데이터를 이용한 감정인식을 병행한다. 이를 통해 스마트폰에서 실용적이며 신뢰도 높은 감정인식 결과를 도출하고 다양한 서비스에서 활용이 가능하도록 함을 목적으로 한다.

II. 관련 연구

1. 컨텍스트 기반 감정인식

스마트폰에서 사용자 선호도와 컨텍스트 정보를 기반으로 사용자의 감정추출 엔진에 대한 연구를 Lee등이 진행하였다[1]. 기존의 스마트폰에서 감정을 인식할 때 사용자에게 요구되는 불편함을 최소화하며 인식한 감정을 기반으로 서비스를 추천하는 애플리케이션을 통해 그 유용성을 평가하였다. 그러나 컨텍스트 정보와 사용자 선호도만으로 추출한 감정은 실제 인식하는 순간의 사용자의 감정과 상이할 수 있다는 단점이 있다.

2. 음성기반 감정인식

A.B Kandali등은 MFCC 특징과 GMM 분류기를 이용하여 음성기반 감정인식 방법을 제안하였다[2]. 음성기반 감정인식은 정확도가 높은 편이지만 인식할 수 있는 감정의 수가 단순하며 짧은 순간의 감정에만 국한되어 있다는 단점이 있다.

III. 본론

1. 시스템 구성

제안하는 시스템의 구조도는 그림 1과 같다.

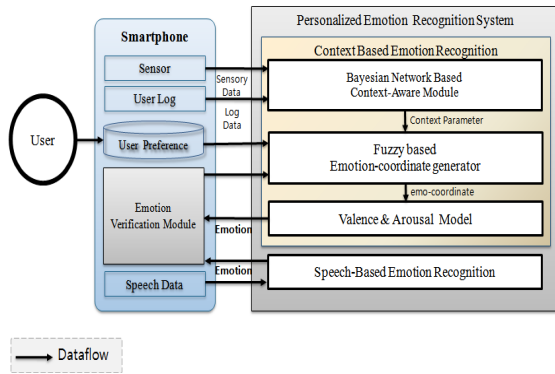


그림 1. 시스템 구조
Fig. 1. System Architecture

스마트폰에서 수집하는 데이터는 사용자 선호도와 사용자 통화, 문자메시지 등의 사용자 로그 및 사용자 통화 음성녹음 데이터이다. 이를 각각 컨텍스트 기반 감정인식과 음성 기반 감정인식 모듈에서 감정을 인식하고 최종적으로 감정 검증(Emotion Verification)모듈에서 비교하여 다음 감정인식에 반영한다.

2. Context Based Emotion Recognition Module

스마트폰에서 수집한 사용자 선호도와 컨텍스트 데이터를 기반으로 베이지안 네트워크를 통해 컨텍스트를 생성하고 이를 퍼지추론과 Valence-Arousal 모델을 활용하여 감정을 인식한다.

3. Speech based Emotion Recognition Module

음성기반 감정인식은 널리 사용되는 MFCC 필터뱅크 알고리즘을 사용한다. 통화 녹음 데이터를 통해 단위 시간 윈도우의 음성 데이터를 MFCC 필터뱅크 알고리즘으로 필터링한 후, 평균, 표준편차, 최대, 최소, 에너지 평균, 에너지 표준편차, 에너지 최대값의

7가지 특징값을 추출한 후 분류기로 SVM을 활용하여 감정을 인식한다.

4. Emotion Verification Module

각각 컨텍스트와 음성 기반으로 인식한 감정은 Emotion Verification Module에서 동일 시간대의 감정을 비교한 후 컨텍스트 기반 감정인식에 반영한다. 이를 통해 각각의 단일 기법에 의한 감정인식의 단점을 상쇄하며 인식한 감정의 신뢰도를 향상시킨다.

IV. 결론

본 논문에서는 스마트폰 환경에서 실용적이면서도 신뢰도 높은 감정인식을 위해 사용자의 선호도와 컨텍스트 정보를 활용하여 감정을 인식하고 음성기반 감정인식을 통해 인식한 감정의 신뢰도를 향상시킬 수 있는 시스템을 제안하였다. 제안하는 시스템을 활용하여 스마트폰에서 감정을 활용한 다양한 서비스 제공이 가능할 것으로 기대된다.

감사의 글

본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 육성지원 사업의 연구결과로 수행되었음 (NIPA-2012-(H0301-12-1004))

참고문헌

[1] Seongho Lee, Jae Hun Bang, Sungyoung Lee, "User Emotion Extraction Engine(E³) based on Fuzzy Inference and Bayesian Networks in Smart Phone Environment," KIISE 2011 KCC Vol. 38, No. 1(A), pp. 100-103, June 2011.
[2] Bihar Kandali, A.; Routray, A.; Kumar Basu, T, "Emotion recognition from Assamese speeches using MFCC features and GMM classifier," TENCON 2008 IEEE Region 10 conference, pp.1-5, 2008