

작품설명서

작품명: 음성 인터페이스를 활용한 감정 인식과 감정 음성 합성에 대한 기술 동향
제출자: 정보과학과 우은주

<연구의 배경>

사회가 발전하고 고도화됨에 따라 사회생활을 하면서 맺는 인간관계 속에서 정신적, 심리적 불안감을 느끼는 사람이 많아지고 있다. 경제발전 및 의료기술의 진전으로 고령 인구의 증가는 노인을 위한 의료 및 사회 복지 서비스라는 사회적 비용의 증가를 초래한다. 또한, 경쟁이 만연한 사회에서 생존을 위한 경쟁 압박은 스트레스, 우울증, 감정표현 불능, 자살 등 다양한 형태로 정서적, 심리적 문제를 발생시키므로 정서적·심리적 불안을 해소할 필요성이 크다.

이러한 사회에서 개인은 자기 자신의 감정을 표현, 감지, 인식하지 못하는 특징이 있어, 자신이 화가 나거나 불안해하는 감정이 정확히 어떤 감정인지 구별하지 못하는 경향이 커진다. 사회 전체적으로 정서적으로 불안, 우울 등의 심리적 상태로 고통받는 개개인의 정신 건강 개선을 위해 감정 치유 기술에 대한 수요는 확대될 전망이다.

생체 신호 인지 기술 및 기계학습, 딥러닝 등의 인공지능 기술의 발전으로 오늘날은 인간의 감정 상태를 파악하는 기술개발이 점차 일상생활에서 적용 가능한 방향으로 연구된다. 그 예로 얼굴의 표정이나 몸짓 등을 파악할 수 있는 영상 인식, 목소리의 강약과 크기 등을 감지할 수 있는 음성 인식, 맥박, 심전도 등의 변화를 감지하는 생체정보 인식 기술을 통해 인간의 내면에 있는 감정 상태 파악이 가능해진 것이다.

이와 같은 과학 기술의 발전으로 사용자의 감정을 읽어내고 이 정보에 기반한 개인화 서비스를 제공하는 기술은 미래의 기술과 산업의 새로운 돌파구를 마련하는 데에 큰 역할을 할 수 있다. 이 분야의 큰 잠재력과 시장성으로 인해 인공지능 및 음성 공학 기술이 발달한 국가들은 감정 인식 분야에 많은 관심을 가지고 지속적인 인공지능 기술개발에 있어서 필수적인 연구주제로 인적·물적 자원을 투자하고 있다.

인간의 감정 상태에 대한 데이터의 증가와 이러한 빅데이터와 인공지능 알고리즘을 활용한 새로운 비즈니스 솔루션 창출 기회의 가능성이 증대되면서 인간의 감성을 반영한 새로운 비즈니스 시장 창출의 기회가 많아짐에 따라 의료, 자동차, 교육, 판매, 광고 등의 다양한 산업과의 융합을 통하여 일상생활에서 사용하는 제품 및 서비스에 인간의 감성을 적용할 수 있다. 일례로 평상시에 소셜네트워크상의 사용 언어, 음성의 미묘한 변화에 대한 감성파악을 통해 우울증

을 진단하고 자살을 사전에 방지하는 등 다양한 서비스 개발이 가능할 것이다.

<연구의 목적 및 방향>

정보·통신기술이 발달한 현대를 살아가는 많은 사람은 정서적, 심리적인 문제를 가지고 있고 이에 대해 사람의 감정을 읽고 이에 대응하는 기술은 육체적, 정신적 상처를 극복하며 살아가야 하는 현대인들에게 매우 중요한 문제이다. 따라서 본 과제는 감성 기반 연구라는 큰 틀에서 음성 인터페이스가 어떻게 사용되고 있는지를 중심으로 진행하였다.

음성 인터페이스는 직관적이고 편리해 키보드나 마우스 같은 별도의 조작이 필요한 장치 없이도 누구나 사용할 수 있는 인터페이스이지만 음성 인식 및 음성 합성은 일반 사람들이 쉽게 이해하기에는 그 기술의 진입 장벽이 높고 그와 관련한 일반적인 교양서도 거의 출판되지 않은 실정이다. 음성은 인간-기계 상호작용에 있어 강력한 도구가 될 수 있으므로 음성 인터페이스에 관한 기술 동향을 감성 인식과 감정이 녹아 있는 음성 합성을 주제로 각각 2장과 3장으로 나누어 정리하였다. 부록으로 감성 음성 합성에 사용되는 W3C 표준 규약인 SSML의 응용 사례를 도입하여 서비스하고 있는 기업별로 정리하였고 SSML의 사용법을 보기 쉽도록 간략하게 서술하였다. 감성 음성기술의 응용 분야로 음성 인터페이스의 로봇공학 분야의 응용 사례를 4장에 정리하였다. 본 보고서에 기재한 난해한 기술 용어는 각주로 정리하였다.