

표정과 음성 감성 분석을 통한 통합 스트레스 분석 시스템

이애진*, 전지원*, 유수화*, 김윤희
 숙명여자대학교 소프트웨어학부

dldowls1211@gmail.com a73277508@gmail.com dhfhfkzjxms@sookmyung.ac.kr
 yulan@sm.ac.kr

An Integrated Stress Analysis System using Facial and Voice Sentiment

Aejin Lee, Jiwon Chun, Suhwa Yu, Yoonhee Kim
 Dept. of Software, Sookmyung Women's University

요 약

현대 사회에서 극심한 스트레스로 고통을 호소하는 사람들이 많아짐에 따라 효과적인 스트레스 측정 시스템의 필요성이 대두되었다. 본 연구에서는 영상 속 인물의 표정과 음성 감성 분석을 통한 통합 스트레스 분석 시스템을 제안한다. 영상 속 인물의 표정과 음성 감성 분석 후 각 감성값에서 스트레스 지수를 도출하고 정량화한다. 표정과 음성 스트레스 지수로 도출된 통합 스트레스 지수가 높을수록 스트레스 강도가 높음을 증명하였다.

1. 서론

코로나 19의 장기화로 인한 외부활동 제한과 경제적 어려움으로 유발된 스트레스가 우울증과 불안을 심화하여 심리·사회적으로 영향을 끼치고 있다. [1] 이는 스트레스가 심각한 사회 문제로 자리 잡았음을 시사한다. 하지만 스트레스를 측정하는 효과적인 방안이 마련되지 않은 실정이다. 기존의 스트레스 측정 시스템인 자가 진단 방법은 주관적이기 때문에 정확성이 낮으며 정밀 검사 방법은 시간적, 비용적 부담이 크다는 문제점이 존재한다. 스트레스는 표정과 음성으로 측정이 가능하므로[2], 따라서 표정과 음성을 분석하여 스트레스를 측정하는 시스템이 필요하다.

본 논문에서는 표정과 음성 감성 분석을 통한 통합 스트레스 분석 시스템(InSAS)을 제안한다. 시스템은 영상 내 인물의 표정 감성값에서 표정 스트레스 지수를 도출, 문장별 감성값에서 음성 스트레스 지수를 도출한다. 이후 통합 스트레스 분석기에서 표정과 음성 스트레스 지수를 이용하여 영상의 통합 스트레스를 측정하는 실험을 진행한다.

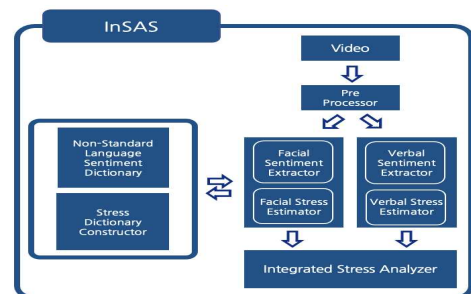
본 논문은 다음과 같이 구성된다. 2장은 관련 연구, 3장은 본 논문에서 제안하는 InSAS의 개요와 기능 설계, 4장에서는 실험 분석 및 결과를 설명하고 5장으로 결론을 맺는다.

2. 관련 연구

[3], [4] 논문에서 anger, sadness, surprise가 스트레스의 주된 감정임을 증명했다. [5] 논문은 가장 오래 지속되는 감정이 sadness이며 surprise, fear가 순간적인 감정임을 밝혔다. [6] 논문에서 감정 노동은 자신의 부정적인 감정을 억누르고 긍정적인 감정을 연기하는 것이며 스트레스 요인이라는 것을 밝혔다.

3. An Integrated Stress Analysis System using Facial and Voice Sentiment (InSAS)

3.1 시스템 구조



(그림 1) InSAS 구조도

InSAS는 영상 속 인물의 표정과 음성을 기반으로 감성 분석을 진행한 뒤 영상의 통합 스트레스를 측정하는 시스템이다. 그림 1은 본 연구의 시스템 구조도이다. 먼저 영상에서 전처리기(Preprocessor)를 통해

영상 구간 프레임 추출 및 이미지 크롭을 진행한다. 표정 감성 추출기(Facial Sentiment Extractor)는 표정의 8가지 감성값을 추출한다. 표정 스트레스 측정기(Facial Stress Estimator)는 문장별 표정 스트레스 지수를 도출한다. 다음으로 구축된 감성 사전[8]에 단어가 없으면 word2Vec[12]을 이용하여 비표준어 감성 사전(Non-Standard Language Sentiment Dictionary)을 구축한다. 음성 감성값 추출기(Verbal Sentiment Value Extractor)는 문장별 Timestamp 기반으로 텍스트를 한 문장씩 분리하고 감성값을 추출한다. 음성 스트레스 측정기(Verbal Stress Estimator)는 문장별 음성 스트레스 지수를 도출한다. 통합 스트레스 분석기(Integrated Stress Analyzer)는 표정 스트레스 지수와 음성 스트레스 지수를 문장 당 0~10 사이의 값으로 나오도록 하고, 상반된 스트레스 지수가 크면 가중치를 주어 0~10 사이로 영상의 통합 스트레스 지수를 도출한다.

3.2 표정 스트레스와 음성 스트레스 지수 도출과 통합 스트레스 지수 측정

표정 스트레스 지수 분석을 위해 Microsoft Azure의 FACE API[13]를 이용하여 표정의 8가지 감성값을 추출한다. 감성이 순간 감성 또는 지속 감성인지 [5]에 따라 감성값 전처리를 진행한 후 문장 단위로 영상 프레임들의 감성값 평균을 도출한다. 이후 스트레스의 주요 감성과 강도를 고려하여 문장 당 0~10 사이의 표정 스트레스 지수를 부여한다.

다음으로 음성 스트레스 지수 분석을 위해 Google Cloud의 Speech-to-Text(STT) API[9]와 Java 기반 형태소 분석기 KOMORAN[10]을 이용하여 텍스트로 변환한 음성 데이터의 형태소 분석을 진행한다. 감성 사전[8]과 비표준어 감성 사전에서 긍정, 부정, 중립의 감성값을 가져온 뒤 이전 연구[11]에서 제시한 언어 감성값 도출 알고리즘을 통해 문장의 감성값을 도출한다. 문장의 감성값에 [7]을 기준으로 정의한 스트레스 단어와 강조 부사어 여부를 판단해 문장별 음성 스트레스 지수를 부여한다. 스트레스 지수는 0일 경우 ‘거의 없음’, 1은 ‘조금 있음’, 2는 ‘상당히 있음’, 3은 ‘매우 심함’으로 나뉜다. 이후 스트레스 통합을 위해 0~10 사이로 값을 수정한다.

표정과 음성 스트레스 지수를 이용하여 영상의 문장별 통합 스트레스 지수를 나타낸다. 음성과 표정의 스트레스 지수가 서로 다를 경우 큰 값을 해당 문장의 스트레스 지수로, 상반 스트레스 지수가 일정 기

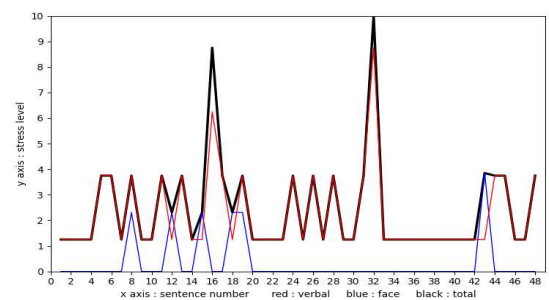
준을 넘을 경우 가중치를 주어 영상의 통합 스트레스 지수를 도출한다. 통합 스트레스 지수가 0~2.5이면 ‘거의 없음’, 2.5~5이면 ‘조금 있음’, 5~7.5이면 ‘상당히 있음’, 7.5~10이면 ‘매우 심함’의 4단계로 나뉜다.

4. 시스템 결과 분석

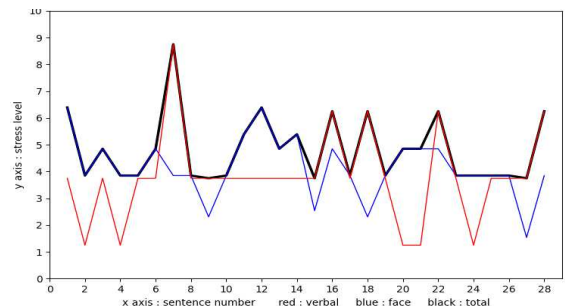
본 연구는 총 6개의 직업별 대표 영상을 선정하여 표정과 음성 스트레스를 실험하였다.

<표 1> 실험 대상 영상

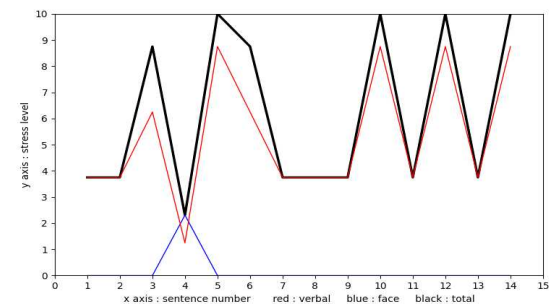
직업군	영상 내용
강형욱 1[14]	에건 훈련사 강형욱의 근황소개
강형욱 2[15]	개물림 사고에 대한 의견
승무원[16]	진상 손님 경험담



(그림 2) ‘강형욱 1’ 스트레스 그래프



(그림 3) ‘강형욱 2’ 스트레스 그래프



(그림 4) ‘승무원’ 스트레스 그래프

그림 2, 3, 4는 각각 ‘강형욱 1’, ‘강형욱 2’, ‘승무원’ 영상의 스트레스 그래프이다. x축은 문장 번호이고 y축은 0에서 10까지의 스트레스 지수이다.

4.1 표정 스트레스 측정

그림5에서 그림10은 ‘강형욱 1’, ‘강형욱 2’, ‘승무원’ 영상의 표정 스트레스 분석 결과 예시이다. 주요

감성값은 표정 스트레스 측정 알고리즘에 의해 전처리된 감성값이다.



(그림 5) '강형욱 1' 표정 스트레스 최소 구간 (그림 6) '강형욱 1' 표정 스트레스 최대 구간

그림 5의 표정 스트레스 지수는 0이다. 그림 6의 주요 감성값으로 sadness와 surprise가 비중 있게 나타났다. sadness가 나타나지만 유의미한 값을 가지지 않고, 스트레스 중심 감성인 surprise에 가중치를 주어 최종 표정 스트레스 지수는 3.85로 측정되었다. 그림 2의 파란 선은 '강형욱 1' 영상의 표정 스트레스 지수 변화 그래프이다. 전반적으로 0에서 2.5 사이의 스트레스를 나타낸다. 이를 통해 '강형욱 1'은 표정 스트레스가 '거의 없음'을 보여준다.



(그림 7) '강형욱 2' 표정 스트레스 최소 구간 (그림 8) '강형욱 2' 표정 스트레스 최대 구간

그림 7의 표정 스트레스 지수는 1.54이다. 그림 8의 주요 감성값으로 fear, sadness, surprise가 비중 있게 나타났다. 부정 감성인 fear와 sadness가 유의미한 값을 가지고, 스트레스의 주된 감성인 surprise가 포함되어 있으므로 가중치를 주어 최종 표정 스트레스 지수는 6.39로 측정되었다. 그림 3의 파란 선은 전반적으로 3.5에서 5.5 사이의 스트레스를 나타낸다. 이를 통해 '강형욱 2'는 표정 스트레스가 '조금 있음'을 보여준다.



(그림 9) '승무원' 표정 스트레스 최소 구간 (그림 10) '승무원' 표정 스트레스 최대 구간

그림 9의 표정 스트레스 지수는 0이다. 그림 10의 주요 감성값으로 happiness와 surprise가 비중 있게 나타났다. 스트레스의 주된 감성인 surprise가 나타나지만, 표정의 긍정 감성값이 유의미하므로 가중치를 받지 않는다. 따라서 최종 표정 스트레스 지수는 2.31이다. 그림 4의 파란 선은 대부분 0이며, 4번째 문장에서 한 차례 2.31의 값이 나타났다. 이를 통해

'승무원'의 표정 스트레스는 '거의 없음'을 보여준다.

4.2 음성 스트레스 측정

<표 2>는 '강형욱 1', '승무원' 영상의 음성 스트레스 분석 결과 예시이다.

<표 2> 문장별 음성 스트레스 지수 예시

영상	번호	문장	감정	스트레스 단어 개수	강조 부사 유무	스트레스 지수
강형욱 1	1	훈련을 조금 더 열심히 하려고 하고 있고 전 이게 행복해서 너무 좋아요	긍정	없음	없음	0단계
	2	내가 하라는 거 제대로 못하고 있어 가지고	부정	1개	있음	3단계
승무원	3	그래서 최대한 이제 눈높이 자세를 취하면서 눈높이 자세란 건 무릎을 살짝 구부리고 손님이랑 눈높이 맞은 자세에서 맞추면서 손님을 컨택을 했던 것 같아요	긍정	없음	없음	0단계
	4	그땐 거 안 먹어 이리면서 오렌지 주스를 탁 치셨는데 주스가 정말 다 난사가 된 거예요	부정	1개	있음	3단계

<표 2>의 1번 문장은 이전 연구[11]의 언어 감성값 도출 알고리즘을 적용하여 긍정 감성값이 부여되었다. 그리고 음성 스트레스 측정 알고리즘에 따라, 스트레스 단어가 포함되지 않아 음성 스트레스 지수는 0단계이다. 이처럼 <표 2>의 4번 문장도 '그땐', '안', '먹-' 등의 부정 감성값을 가진 형태소로 인해 부정 감성값이 부여되었으며, 스트레스 측정 알고리즘에 따라 스트레스 단어인 '치-'와 강조 부사인 '정말'이 존재하므로 음성 스트레스 지수는 3단계이다.

그림 2의 빨간 선은 '강형욱 1' 영상의 음성 스트레스 지수 변화 그래프이다. 전반적으로 1에서 4의 값을 보이나, 16번째 문장에서 6.25와 32번째 문장에서 8.75의 특징적인 값이 나타난다. 해당 문장은 '꿈인데 안 돼도 할 수 없죠'와 '내가 하라는 거 제대로 못 하고 있어 가지고'로 근황 소개 중 훈련에 대해 아쉬움을 토로하며 높은 스트레스 지수를 기록했으나 영상 자체는 음성 스트레스가 높아 보이지 않는다.

그림 3의 빨간 선은 전반적으로 2.5에서 6 사이의 스트레스이므로 '강형욱 2'의 음성 스트레스는 '조금 있음'과 '상당히 있음'을 아우름을 보여준다.

그림 4의 빨간 선은 전반적으로 4에서 8 정도의 스트레스이므로 '승무원'의 음성 스트레스는 '조금 있음'과 '매우 심함'을 아우름을 보여준다.

4.3 통합 스트레스 분석

그림 2, 3, 4의 감정 선은 각각 ‘강형욱 1’, ‘강형욱 2’, ‘승무원’ 영상의 통합 스트레스 그래프이다.

그림 2는 48개의 문장 중 과반수 이상인 28개가 ‘거의 없음’ 단계에 속한다. 반면 그림 3은 28개의 문장 중 과반수 이상인 16개가 ‘조금 있음’ 단계이고 11개가 ‘상당히 있음’ 단계이다.

그림 4에서 통합 스트레스가 가장 높았던 부분의 문장인 ‘그래서 진심을 다해 서비스를 해도 이미 고객한테 실수한 것은 통하지 않는구나 이런 생각이 들어서 되게 마음이 아팠죠’는 표정 스트레스가 0이고, 음성 스트레스가 3단계이다. 해당 문장은 표정 스트레스와 음성 스트레스에서 상당한 차이를 보인다. 이처럼 두 스트레스의 차이가 크면 자신의 부정적인 감정을 억누르고 긍정적인 감정을 연기하는 것이며 스트레스 요인이라는 연구 결과[6]에 따라 통합 스트레스 지수에 가중치를 준다. 따라서 이 문장의 통합 스트레스 지수는 8.75이고, 승무원이 감정 노동자로서 매우 심한 스트레스를 받는다고 판단할 수 있다.

5. 결론

본 연구에서는 영상 속 인물의 표정과 음성을 통해 표정 감성값과 음성 감성값을 추출하였다. 추출된 감성값들을 바탕으로 표정과 음성의 스트레스 지수를 도출하였으며, 이들을 통합해 영상에서 문장별 최종 스트레스 지수를 객관적이고 정량적으로 제시하였다. 향후 연구는 감정노동자뿐만 아니라 다양한 직업군에 대한 스트레스 분석을 진행할 것이다. 또한, 비표준어 감성 사전을 확대 구축하고 문장 음성 감성값 도출 알고리즘을 발전시킴에 따라 음성 스트레스 지수 도출에 정확도를 가할 수 있다.

이 논문(저서)은 2021년도 정부(교육부)의 재원으로 한국과학창의재단(2021년도 학부생 연구프로그램)의 지원을 받아 수행된 연구임.

참고문헌

[1] Gyeong Dae Ha, “Half of the people experienced depression and anxiety due to COVID-19…The biggest reason is poor sleep quality”, medigatenews, 2021.7.29.
 [2] 장은혜, 박병준, 김상협, 손진훈, “생체신호 기반의 부적 감성 분류를 위한 최적 기계 학습 알고리즘 검증.” 한국감성과학회 2011년도 추계학술대회, 경

주, 한국, 2011, pp.63-64.

[3] Richard S. Lazarus, “Stress and Emotion : A New Synthesis,” New York, Springer Publishing Company, pp.35-38, 2006.
 [4] Daudelin-Peltier, C., Forget, H., Blais, C. et al, “The effect of acute social stress on the recognition of facial expression of emotions,” Sci Rep 7, 1036, 2017.
 [5] Verduyn, P., Lavrijsen, S., “Which emotions last longest and why: The role of event importance and rumination,” Motiv Emot 39, pp.119 - 127, 2014.
 [6] Rehman, Saqib & Hamza, Muhammad & Anum, Leena & Zaid, Farah & Khan, Ahmed hussain & Farooq, Zahida, “What if it is too negative? Managing emotions in the organization,” Management Science Letters, pp.1489-1498, 2021.
 [7] Sun Ju Sohn, Mi Sook Park, Jieun Park, Jin Hun Sohn, “Korean Emotion Vocabulary: Extraction and Categorization of Feeling Words,” Science of Emotion & Sensibility, 15(1), pp.105-120, 2012.
 [8]<https://github.com/park1200656/KnuSentiLex>
 [9] Google Cloud의 Speech-to-Text(STT) API <https://cloud.google.com/speech-to-text?hl=ko>
 [10] JAVA 기반 형태소 분석기 KOMORAN <https://github.com/shineware/KOMORAN>
 [11] 신희원, 이소정, 손규진, 김혜린,곽서현, 김영민, 김윤희, “언어와 비언어 표현의 통합 분석을 통한 리뷰 동영상의 신뢰성 분석,” 한국정보처리학회 학술대회논문집, Vol.27, No.2, pp.515-518, 2020.
 [12] Gensim, Word2Vec [Internet], <https://radimrehurek.com/gensim/>
 [13] FaceAPI, Microsoft Azure [Internet], <https://www.azure.microsoft.com/en-us/services/cognitive-services/face/>
 [14]<https://www.youtube.com/watch?v=woBQb-wJr8o&list=PLVh3TM0B0Wtn0TYSjIPcyY7VSahjCgnXy>
 [15]https://www.youtube.com/watch?v=K_4TPvXLQso
 [16]<https://www.youtube.com/watch?v=ea54mKRD2Ks>