

라즈베리 파이 기반의 음성 감정 분석 시스템을 통한 감성적 이미지 색상 전달

김종현^o

^o강남대학교 소프트웨어응용학부

e-mail: jonghyunkim@kangnam.ac.kr^o

Emotional Image Color Transfer via Voice Emotion Analytics System Based on Raspberry Pi

Jong-Hyun Kim^o

^oDept. of Software Application, Kangnam University

● 요약 ●

본 논문은 일상적인 대화로부터 감성을 추출하고 분석함으로써 상황에 맞는 대화의 내용과 분위기를 이미지의 색상으로 표현할 수 있는 이미지 색상 변환 프레임워크를 소개한다. 본 연구는 라즈베리 파이와 마이크 센서를 기반으로 사용자로부터 목소리를 입력받을 수 있는 모듈을 제작하고, 그 목소리로부터 감성을 분석한다. 분석된 감성을 이용하여 이미지의 색상을 자동으로 변환하는 기술과 통합함으로써 청각장애인 및 미취학 아동들이 화자의 대화를 이미지를 통해 쉽게 인지하여 의사소통 및 감성 전달 환경을 개선하고자 한다.

키워드: 이미지 색상 전달(Image color transfer), 음성 감정 분석 시스템(Voice emotion analytics system), 라즈베리 파이(Raspberry pi)

I. Introduction

이미지 색상 변환은 영상처리와 컴퓨터 그래픽스 분야에서 폭 넓게 사용되고 있다[1,2]. 본 연구에서는 이미지의 색상에 사용자 감성을 넣어 전달하는 방식을 기반으로 청각적 신호를 정확하게 전달받지 못하는 청각장애아동과 노인들을 위한 감성적 이미지 색상 프레임워크를 제안한다.

청력 손실을 보상해주는 보청기, 수화, 구화, 필담 등의 기법들이 존재하지만 주위의 소음이나 말소리에 해당되는 주파수 일부가 왜곡되어 의사소통이나 학습에 방해를 받는다[3]. 같은 대화에서도 화자의 상황이나 감정에 따라 문장은 다르게 해석된다. 그 형태는 주파수로도 쉽게 확인되며[4] 이러한 차이가 실제 의사소통에 문제가 될 수 있다는 것을 발견하여 본 논문에서는 이 문제를 최소화하기 위한 감성적 이미지 색상 변환 솔루션을 제안한다.

II. The Proposed Scheme

본 논문의 알고리즘 개요는 일반적인 대화를 입력으로 아래와 같은 순서로 수행된다 (Fig. 1 참조).

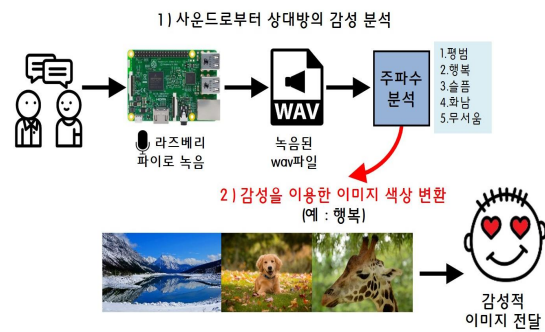


Fig. 1. Algorithm overview.

1. 라즈베리 파이에 USB 마이크 센서를 연결하고 이를 통해 대화를 자동으로 녹음
2. 녹음된 대화의 Wav파일을 PC의 로컬 메모리에 저장하고 이를 자동으로 읽어 들임
3. Wav파일로부터 감성을 분석
4. 분석된 감성과 연관되는 분위기 색상을 결정하고 이를 기반으로 이미지의 색상을 변환

기존 접근법과의 차이는 사용자의 목소리로부터 감성을 분석하고, 분석된 사용자의 감성을 이미지의 색상 변환을 통해 사용자의 감성을 표현한다는 점이다.

1. Voice Input Module using Raspberry Pi

본 연구에서는 USB 마이크 센서를 이용하여 라즈베리 파이에서 사운드 녹음 기능이 실행되도록 디자인했으며, 원격 터미널 작업을 통해 사운드 파일을 저장하였다. 화자들의 대화가 사운드 센서를 통해 녹음되면, 터미널을 통해 녹음된 Wav파일을 저장한다 (Fig. 2 참조).

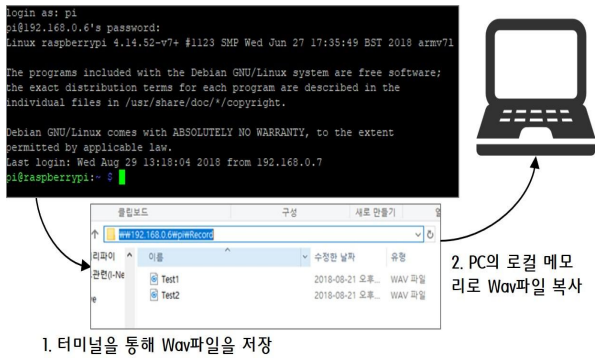


Fig. 2. Saving voice's wav file using terminal.

Fig. 2에서 보듯이 터미널을 통해 녹음된 Wav파일은 외부 원격 메모리에 저장되며, 터미널 명령어를 이용하여 이 Wav파일을 PC의 로컬 메모리로 저장한다. 로컬 메모리에 Wav파일이 저장되면, 분석 프로그램에 의해 자동으로 Wav파일의 감성이 분석되며 최종적으로 이미지 색상 변환과정이 수행된다.

2. Emotion Analysis from Sound and Image Color Transfer

본 연구에서는 5개의 감성 카테고리(평범, 행복, 슬픔, 화남, 무서움)를 이용하여 사운드로부터 감성을 추출한다 (Fig 3 참조). 사운드로부터 감성을 추출하기 위해 5가지 감성의 평균과 표준편차를 이용하여 감성을 분류하였으며, 이를 구현하기 위해 *OpenVokaturi* 라이브러리를 이용하였다.

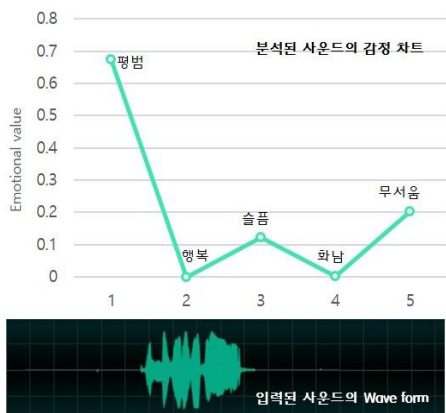


Fig. 3. Emotional chart analyzed from input sound.

입력 이미지의 색상을 변환하기 위해 본 논문에서는 목소리로부터 계산된 감성 값을 기반으로 연관되는 색상을 선택한 후, 다양한 이미지

색상 변환 방법 중 Reinhard 등의[5] 방법을 사용해 색상을 변환하였으며, 감정에 연관되는 색상을 결정하기 위해 Yang과 Peng[6]에서 사용한 24개의 감정에 대한 색 조합을 사용하였다.



Fig. 4. Emotion templates.

본 연구에서는 사운드의 5개의 감정 중 화남과 무서움은 부정의 의미가 강한 비슷한 감정이므로 하나로 묶었다. 결과적으로 4개의 감성과 색 조합을 연결시킬 때는 일 대 다 관계로 설정하여 하나의 감정이 유사한 다수 개의 감정 색을 갖도록 연결시켰다. 각 감정 색에 해당하는 참조 이미지들은 Fig. 4와 같으며, 우리는 각 감정에 연관된 참조 이미지들을 이용하여 이미지 색상 변환을 수행한다.

III. Results

본 논문에서는 사운드의 감성을 이용하여 의사소통에 도움을 줄 수 있는 이미지 시각화 시스템을 제안했다. 우리는 일상적인 대화 템플릿을 이용하여 실험했으며 (Table 1 참조), 대화의 감성 인식 측면에서 유의미한 결과를 얻었다.

Table 1. Conversation templates.

지금 몇 시야, 지금 뭐해, 빨리 와, 빨리 왔네, 이게 뭐야, 잘했네, 등

Fig. 5는 본 논문에서 화자의 대화를 입력 받을 수 있도록 설계한 입력 시스템이며, USB 마이크를 이용하여 사용자가 쉽게 조작이 가능하도록 제작하였다. 사용자는 USB 마이크에 근접하여 말을 하면 음성 파일이 녹음되어 Wav파일로 저장된다. 이 Wav파일을 라즈베리 파이에 연결된 노트북의 로컬 메모리에 저장하고 분석 프로그램에 의해 감성을 분석하고 이미지 색상 변환을 수행한다.

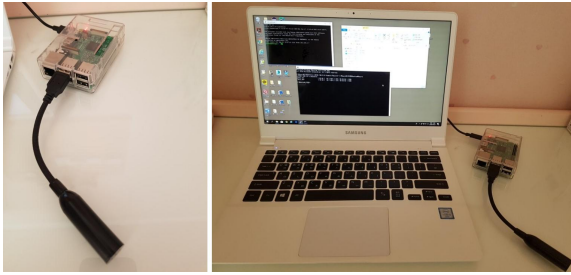


Fig. 5. Input module system using raspberry pi and USB microphone sensor.

Fig. 6은 분석된 감성을 기반으로 이미지 색상을 변환한 결과이다. USB 마이크로 입력된 사용자의 목소리의 감정에 따라 색상이 변환된 결과를 보여주고 있으며, 각 감성에 대한 분류된 이미지 템플릿(Fig. 4 참조)을 참조하여 변환되기 때문에 사용자의 감성을 잘 표현한 결과를 보여준다.

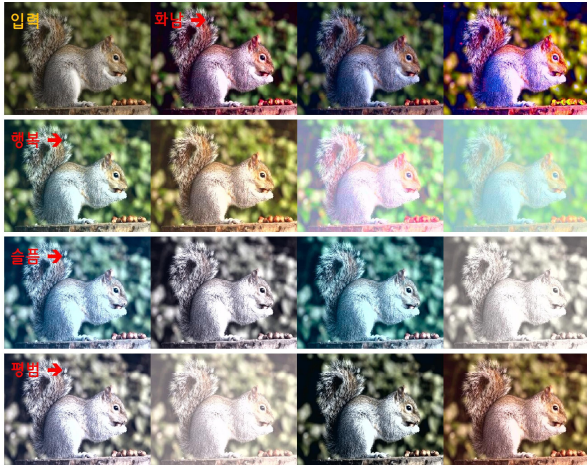


Fig. 6. Image color transfer with our method.

REFERENCES

- [1] Reinhard, Erik, Michael Adhikhmin, Bruce Gooch, and Peter Shirley. "Color transfer between images." IEEE Computer Graphics and Applications, pp. 34-41, 2001.
- [2] Cao, K., Liao, J., & Yuan, L., "CariGANs: Unpaired photo-to-caricature translation", ACM SIGGRAPH 2018, arXiv preprint arXiv:1811.00222.
- [3] Kyoungghan Park, "A study on the characteristics of connective ending use by connective ending forms for the students with hearing impairment", The Korean Society of Education for Hearing-Language Impairments, 2014, pp. 1-22.
- [4] SooHo Lee, ChaeYoung Shin, YoungChan Shin, SeungYong Ji, JiHye Choi, DaeWoong Bae, YunJin Jo,

- JiWoo Kim, SeYoung Shin and Jong-Hyun Kim, "Communications helper with emotional filtering of voice," Proceedings of the Korea Computer Graphics Society Conference 2018, pp. 45-46.
- [5] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley, "Color transfer between images", IEEE Computer Graphics and Applications 2001, pp. 34-41.
- [6] Chuan-Kai Yang, Li-Kai Peng, "Automatic mood-transferring between color images", IEEE Computer Graphics and Applications 2008, pp. 52-61.