



# 시각장애인을 위한 음성 도우미 장치

채준기\*<sup>1</sup>, 장지우\*<sup>2</sup>, 김동완\*<sup>3</sup>, 정수진\*<sup>4</sup>, 이익현\*\*

## Voice Assistant for Visually Impaired People

Jun-Gy Chae\*<sup>1</sup>, Ji-Woo Jang\*<sup>2</sup>, Dong-Wan Kim\*<sup>3</sup>, Su-Jin Jung\*<sup>4</sup>, and Ik Hyun Lee\*\*

이 논문은 2018년도 한국산업기술대학교 학술연구진흥사업에 의하여 연구되었으며, 정부의 재원으로 한국연구재단의 지원을 받아 수행된 사업임(No. 2017R1D1A1B03033526, NRF-2017R1A6A1A03015562)

### 요약

시각장애인은 일상생활에서 시각저하로 인하여 색상, 지폐 금액 확인, 온도 판별 등의 불편을 겪는다. 그리하여 시각장애인을 도와주기 위한, 광학 카메라와 열화상 카메라를 이용한 시스템을 제안한다. 광학카메라는 특징 추출을 이용하여 색상 판별, 지폐 금액 확인하며, 온도 판별을 위해서 열화상 카메라를 이용한다. 버튼을 이용하여 색상, 지폐, 온도 모드를 선택하면, 해당 정보를 스피커를 통하여 음성으로 알려준다. 색상은 16가지, 지폐는 4가지, 온도는 4단계로 구분할 수 있으며, 각기 정확도는 90%이다. 향후 입력 영상 블록화, 머신러닝, 열화상 카메라 상위 버전을 통해 정확도를 개선할 수 있다. 또한, 휴대하기 편하게 지팡이에 부착하여 시각 장애인들이 더 편리하게 사용 할 수 있도록 개선 예정이다.

### Abstract

People with compromised visual ability suffer from many inconveniences in daily life, such as distinguishing colors, identifying currency notes and realizing the atmospheric temperature. Therefore, to assist the visually impaired people, we propose a system by utilizing optical and infrared cameras. In the proposed system, an optical camera is used to collect features related to colors and currency notes while an infrared camera is utilized to get temperature information. The user is enabled to select the desired service by pushing the button and the appreciate voice information are provided through the speaker. The device can distinguish 16 kinds of colors, four different currency notes, and temperature information in four steps and the current accuracy is around 90%. It can be improved further through block-wise input image, machine learning, and a higher version of the infrared camera. In addition, it will be attached to the stick for easy carrying and to use it more conveniently.

### Keywords

visually impaired people, optical camera, infrared camera

\* 한국산업기술대학교 메카트로닉스공학과  
 - ORCID<sup>1</sup>: <https://orcid.org/0000-0002-1193-7443>  
 - ORCID<sup>2</sup>: <https://orcid.org/0000-0002-2963-6526>  
 - ORCID<sup>3</sup>: <https://orcid.org/0000-0002-4303-146X>  
 - ORCID<sup>4</sup>: <https://orcid.org/0000-0002-6018-621X>  
 \*\* 한국산업기술대학교 메카트로닉스공학과 교수  
 (교신저자)  
 - ORCID: <https://orcid.org/0000-0002-0605-7572>

• Received: Jan. 07, 2019, Revised: Apr. 17, 2019, Accepted: Apr. 20, 2019  
 • Corresponding Author: Ik Hyun Lee  
 Dept. of Mechatronics, Korea Polytechnic University, 237  
 Sangdaechak-ro, Shiheung-si, Gyeonggi-do, 15073 Korea  
 Tel.: +82-31-8041-0467, Email: ihlee@kpu.ac.kr

## 1. 서 론

대한민국 등록 장애인 수는 270만 명이며, 이 중 시각장애인은 25만 명이다. 이 중 점자해독이 가능한 시각장애인은 6.1%에 불과하다[1]. 시각장애인을 위한 보조기기 대부분은 점자를 출력하는 수백만 원을 호가하는 고가의 기기이며, 이는 장애인의 소득과 생활여건을 고려하였을 때, 개인이 구비하기가 매우 어렵다. 몇몇 기관에 배치되어 사용가능한 하지만 외출을 두려워하는 것과 일상생활을 고려한 기기는 현저히 부족한 현실이다.

시각장애인은 일상생활에 접하게 되는 다양한 물체에 대해 구체적인 정보(색상, 지폐, 온도 등)를 파악하는데 어려움을 겪는다[2]. 시각장애인이 원하는 색상의 옷을 구매시 도우미 없이는 혼자서 색상 파악이 불가능하다. 또한 계산할 때 지폐의 금액을 알아야 하는데, 노후화된 지폐는 점자가 닳아 없어지기 때문에 촉각으로 지폐의 종류(금액)를 구분하기가 어렵다. 가장 심각한 시각장애인 상해사고 1위가 화상 사고인데, 비장애인의 경우 김이나 끓는 모습으로 구분 가능하지만, 시각장애인들은 음식의 온도를 알지 못하기 때문에 쉽게 데이는 경우(손, 입등)가 많다.

그리하여 본 논문에서는 시각장애인의 일상생활에 필요한 정보를 음성으로 전달해주는 시스템을 제안하고자 한다. 이 장치는 크게 물체를 촬영하는 카메라(광학: 색상, 지폐 판별, 열화상: 온도 판별)와

영상으로부터 물체에 관한 정보를 음성으로 출력하는 부분으로 구성된다. 또한 시각장애인이 장치를 손쉽게 편리하게 사용하기 위해서는 한 손으로 조작이 가능하여야 하고, 사용법이 간단해야 한다. 또한 일상생활에서 주로 사용되기 때문에 휴대가 가능해야 하며, 생활여건과 정부지원을 고려하여 저렴해야 한다. 그리하여 휴대 가능성을 고려하여 한 손으로 들 수 있는 크기로 제작하며, 모양이 한 손으로 잡았을 때 편하도록 중간에 굴곡을 만들어 준다. 시각장애인은 자신이 무엇을 촬영하고 있는지 확인하기 어렵기 때문에 장치의 뒷면에 카메라를 장착하며, 광학 카메라와 열화상카메라는 세로로 배치하면 카메라를 손으로 가릴 수 있으므로 굴곡 위 부분에 최대한 가깝게 두 개의 카메라를 설치한다. 알고리즘에서는 16가지 색상 판별, 4가지의 지폐 판별, 4단계의 온도를 구분하도록 설계한다.

## II. 음성 도우미 장치의 특성

음성 도우미 장치는 하드웨어(버튼, 스피커, 단자, 카메라)와 소프트웨어(색상, 지폐, 온도)로 구분된다.

### 2.1 하드웨어

그림 1은 시각장애인을 위한 음성 도우미 장치의 외형도이다. 장애인들이 사용하는 장치는 다루기 쉬어야 하며, 조작법이 간단하여야 한다.

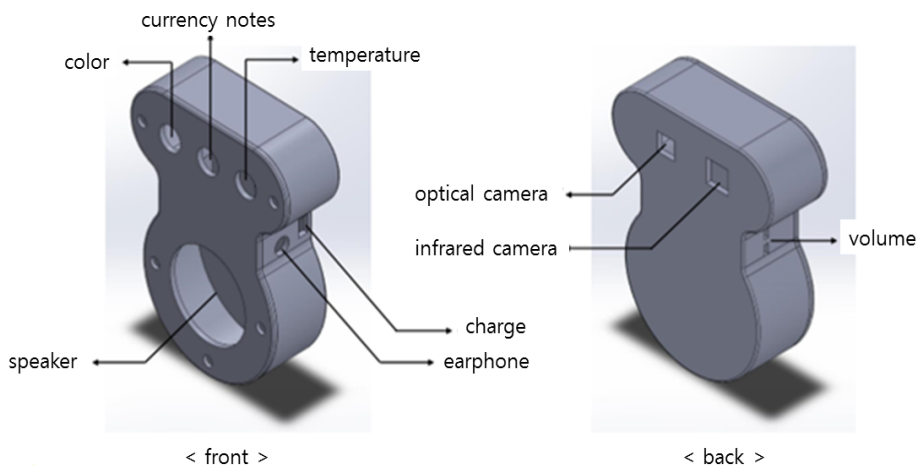


그림 1. 음성 도우미 장치 외형도  
Fig. 1. Prototype of voice assistant

그리하여 전면에 색상, 지폐, 온도를 선택할 수 있는 버튼으로 구성하였으며, 후면에는 광학 카메라와 열화상 카메라로 구성하였다. 또한 우측에는 소음에도 잘 들을 수 있도록 이어폰을 연결할 수 있도록 단자를 만들었으며, 부담 없이 휴대하기 위하여 한손 크기로 제작하였으며, 충전할 수 있는 단자로 구성하였다. 또한 좌측에는 볼륨 스위치를 배치하여 음량을 조절 가능하게 하였다. 출력은 스피커를 통해서 음성으로 알려주며, 이어폰으로 들을 수 있게 이어폰 단자를 부착한다. MCU(Micro Controller Unit)은 영상 처리에 적합한 라즈베리파이 3를 선정하며, 연동 가능한 일반 카메라와 가격이 상대적으로 저렴한 80x60 화소의 열화상 카메라 버전 1을 선택하였다.

## 2.2 소프트웨어

### 2.2.1 색상 판별

색상 판별을 위하여 라즈베리파이용[3][4] 광학카메라를 이용하여 영상을 획득한다. 본 기능은 시각장애인이 옷 색상을 쉽게 파악할 수 있도록 도와준다. 세 가지 모드에서 색상 버튼을 누르면 영상이 획득되며, RGB 영상을 식 (1)을 이용하여 색상-채도-명도 색 공간(HSV, Hue-Saturation-Value)[5]으로 매칭시켜 해당 색상에 대해 영상을 그림 2와 같이 이진화한다. 변환시킨 후 카메라 중앙 영역에서 획득된 부분 중 가장 큰 값을 대표 색상값으로 사용한다.

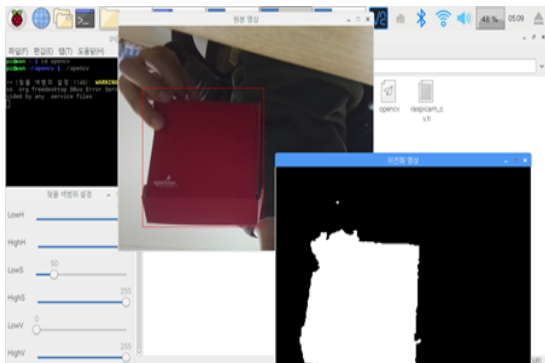


그림 2. 색상 판별을 위한 색상-채도-명도(HSV) 값 매칭  
Fig. 2. HSV matching for distinguishing color

$$S = \begin{cases} \frac{V - \min(R, G, B)}{V}, & \text{if } V \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$H = \begin{cases} 60(G - B) / (V - \min(R, G, B)) & \text{if } V = R \\ 120 + 60(B - R) / (V - \min(R, G, B)), & \text{if } V = G \\ 240 + 60(R - G) / (V - \min(R, G, B)), & \text{if } V = B \end{cases}$$

$$V = \max(R, G, B)$$

### 2.2.2 지폐 판별

본 기능은 지폐의 종류를 음성으로 알려줌으로써, 시각장애인이 쉽게 금액을 파악할 수 있도록 도와준다. 특징 추출과 매칭은 opencv의 특징 매칭 함수를 이용하였으며, 각 특징의 거리가 어느 정도 일 때 두 이미지의 유사도를 판단한다. 전체 지폐의 40%만으로도 구분 가능하며 접힘, 구김, 좌우반전, 상하반전 영상이라도 구별 가능하다. 그림 3은 실제 지폐 판별을 위한 특징 추출과 매칭을 보여준다.



그림 3. 지폐 판별을 위한 특징 매칭

Fig. 3. Feature matching for identifying currency notes

지폐 버튼을 누르면 장치에 저장해둔 지폐(천원, 오천원, 만원, 오만원)가 특징 추출 및 매칭 함수를 이용해 지폐를 비교한다. 천원(F1000), 오천원(F5000), 만원(F10000), 오만원(F50000)을 각 지폐별로 특징 매칭을 하고 카운트 한 값 중 Fmax를 가장 큰 값이라고 정의하고 Fmin을 가장 작은 값이라고 정의한다.

- (1)  $F_{\max}(\text{특징 매칭된 값}) < T_1$  (실험적으로  $T_1$  값을 8로 정의) : 특징 매칭된 값이 T 보다 작으면, 지폐가 없거나 검출이 잘 되지 않았다고 판단한다.
- (2)  $F_{\max}(\text{특징 매칭된 값}) \geq T_1$  : 특징 매칭된 값이 T 보다 크거나 같으면, 지폐가 검출되었다고 판단하여 가장 큰 값을 대표 지폐 금액으로 판단한다.
- (3)  $F_{\min}(\text{특징 매칭된 값}) > T_2$  (실험적으로  $T_2$  값을 40으로 정의) : 특징 매칭된 값이 40보다 크면,

화면에 지폐가 한 장이 아닌 여러 장이 있는 걸로 판단한다. 현재는 가장 큰 값을 대표 지폐 금액으로 판단한다.

2.2.3 온도 판별

본 기능은 시각장애인으로 하여금 뜨거운 물체로부터 화상을 방지할 수 있도록 도와준다. 온도 판별은 flir lepton version 1 열화상 카메라[6]를 이용하였으며, 카메라 내부온도는 식 (2)를 이용한다.

$$\text{온도}(K) = 8,192 - \left( \text{내부온도} \times \left( \frac{\text{테스트프레임} - 8,102}{\text{실험온도} - \text{내부온도}} \right) \right) \quad (2)$$

온도 판별 버튼을 누르게 되면 열화상 영상을 획득하고, 영상 프레임과 온도함수를 이용하여 온도를 구한다. 각 화소의 온도 값을 저장하고 가장 큰 값을 그림 4와 같이 temp(max)로 활용하여 음성 출력한다.



그림 4. 온도 판별을 위한 매칭  
Fig. 4. Realizing the atmospheric temperature

열화상 카메라 측에서 준 내부 온도와 출력 영상의 관계를 통해 각 화소에 온도를 파악한다. 온도의 범위는 “괜찮습니다.”, “약간 뜨겁습니다.”, “뜨거우니 조심 하세요.”, “매우 뜨거우니 조심하세요.” 구분하였다.

2.2.4 음성부 스레드 구현

사진을 찍고 판단하여 음성을 출력하는 것을 순차적으로 진행하게 되면 음성을 말하는 순간 CPU는 동작 하지 않았다. 이런 교착 상태를 방지하기 위해 촬영순간 음성을 출력하고 다음 동작을 예약할 수 있게 스레드 동작을 사용하였다.

표 1은 색상, 지폐, 온도 판별 스레드 전 동작과 후 동작에 따른 지연시간을 나타낸다.

표 1. 음성부 스레드 데이터

Table 1. Thread data for audio part

	before	after
color	within 1 second	within 1 second
currency notes	6.1 seconds	5.2 seconds
temperature	within 1 second	within 1 second

2.2.5 듀얼 터미널 사용을 위한 슬립 구현

색상과 지폐는 opencv[7]-[9]를 사용하고 온도모드에는 qt를 사용한다. opencv와 qt는 화면을 구성하는 라이브러리로 두 개의 파일을 혼용할 수 문제점이 있다. 두 개의 터미널을 사용하였지만 두 개의 터미널을 사용하면 동작속도가 두 배가 되었다. 그리하여 색상판별 버튼과 지폐판별 버튼이 눌리게 되면 음성이 출력되는 동안 온도 프로그램을 슬립(Sleep)시키고 온도판별 버튼을 누르면 색상판별, 지폐판별 프로그램을 슬립 시켰다.

표 2는 색상, 지폐, 온도 판별 슬립 전 동작과 후 동작에 따른 지연시간을 표로 나타내었다.

표 2. 슬립 사용에 따른 지연시간 데이터

Table 2. Time lag data for using sleep

	before	after
color	1.8 seconds	within 1 second
currency notes	10.5 seconds	5.2 seconds
temperature	1.4 seconds	within 1 second

III. 결 과

3.1 색상 판별

라즈베리파이3와 광학카메라를 사용하여 영상을 획득한다. 기존 색상에 대한 색상-채도-명도(HSV) 값을 라즈베리파이3에 미리 저장해 두고 획득한 영상의 색상-채도-명도(HSV) 값을 비교하여 그림 5와 같이 총 16가지 색상(빨간색, 초록색, 노란색, 살구색, 연두색, 분홍색, 주황색, 검정색, 파란색, 자주색, 하늘색, 연한 보라색, 연한 분홍색, 옥색, 회색, 흰색)을 스피커로 음성 출력하였다.

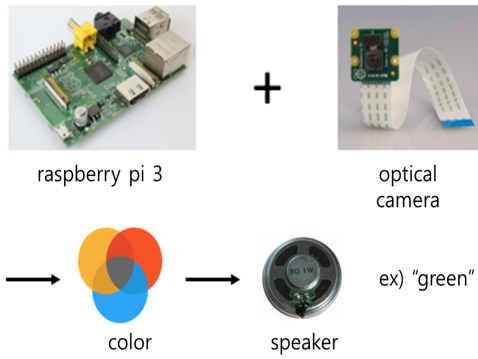


그림 5. 색상 판별

Fig. 5. Color distinguishing

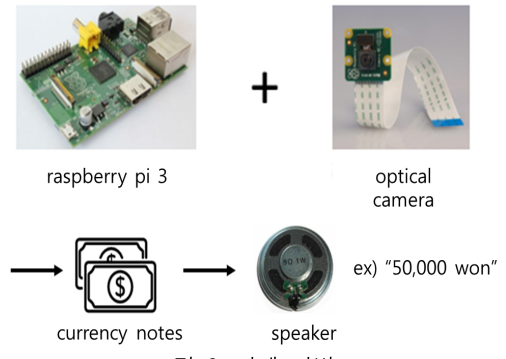


그림 6. 지폐 판별

Fig. 6. Currency notes identification

### 3.2 지폐 판별

라즈베리파이3와 광학카메라를 사용하여 영상을 획득한다. 기존 지폐에 대한 이미지를 라즈베리파이 3에 미리 저장해 두고 획득한 영상과 기존 영상을 특징 추출 및 매칭을 통해 그림 6과 같이 천원, 오천원, 만원, 오만원으로 음성으로 출력하였다. 이때 사용되는 특징 매칭은 각 특징의 거리가 어느 정도 일 때, 두 영상의 유사도를 판단한다. 상기 프로그램은 10 이하일 때를 유사하다고 판단하는 것으로 설정하였다.

작성한 알고리즘에 따라 유사한 것으로 판단된 특징의 개수를 리턴 하였고, 메인함수에서는 유사한 특징이 한 개 이상이면 동일한 영상이라고 판단하도록 구성하였다. 이를 통해 지폐가 구겨지거나 반으로 접어도 획득한 영상과 저장해둔 영상을 비교하여 지폐를 구분하고 음성출력 하도록 하였다.

### 3.3 온도 판별

영상은 그림 7과 같이 라즈베리파이3와 열화상 카메라를 사용하여 획득한다. 900~14,000 나노미터의 영역을 통해 사물의 형상을 인식하는 Flir Lepton Dev Kit 열화상 카메라를 라즈베리파이3 GPIO 핀들에서 I2C와 SPI를 사용하여 연결하였다. 온도 범위에 따라 40도 이하는 “괜찮습니다.” 40도에서 60도 사이는 “약간 뜨겁습니다.” 60도에서 80도 사이는 “뜨거우니 조심하세요.” 80도 이상은 “매우 뜨거우니 조심하세요.”로 음성 출력하였다.

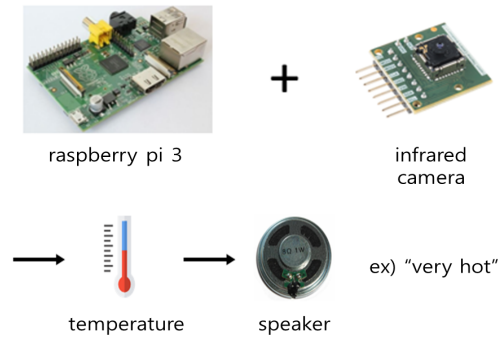


그림 7. 온도 판별

Fig. 7. Atmospheric temperature realization

## IV. 결 론

본 논문에서는 시각장애인의 일상생활의 불편함을 해소하기 위하여 색상, 화폐, 온도 판별 세 가지 기능을 제안하였으며, 각 기능의 정확도는 90%이다. 현재 현재 색상 판별은 중심 영역에 대해서만 판별을 하는데, 입력 영상을 영역(3등분, 9분등)을 지정하면 중앙 영역의 색상 뿐만 아니라 주변 영역의 색상 또한 판별이 가능하다. 또한 현재 16가지 색상을 구분하나, 알고리즘 개선을 통해 밝은 색, 어두운 색 등을 추가 할 수 있다. 지폐 모드는 4가지 지폐판별은 처리 시간이 다소 걸리는 것을 감안하여 복잡한 알고리즘보다 간단한 머신러닝 기법을 사용하여 처리 속도를 개선할 수 있다. 또한 현재 전체 지폐의 40%만으로도 구분 가능하여 전체 지폐의 40% 이하인 경우에는 정확도가 낮다. 그리하여 색상 판별 알고리즘을 적용하면 적은 영역에 대해서도 지폐의 생각을 이용하여 지폐 판별의 정확도를 높일 수 있다. 온도 판별은 현재 Flir lepton 버전 1을 사용하는데, 버전 3.5를 사용하여 제공된 보

정값을 이용하면 정확도를 높일 수 있다. 또한 현재는 온도만 알려주지만 색상과 더불어서 음식에 대한 정보를 알려주는 것으로 개선 할 수 있다. 또한 휴대하기 편하게 목걸이 형태와 시각장애인들의 지팡이에 부착하여 시각 장애인들이 더 편리하게 사용할 수 있도록 개선 예정이다.

### References

- [1] A survey of the visually impaired people by Ministry of Health and Welfare, 2015.
- [2] S. Cardin, D. Thalmann, and F. Vexo, "A wearable system for mobility improvement of visually impaired people", *The Visual Computer*, Vol. 23, No. 2, pp. 109-118, Jan. 2007.
- [3] Ashwin Pajankar, "Raspberry Pi Computer Programming", Packt Publishing, 2015.
- [4] Rushi Gajjar, "Raspberry Pi Sensors", Packt Publinsing, 2015.
- [5] V. Chernov, J. Alander, and V. Bochko, "Integer-based accurate conversion between RGB and HSV color spaces", *Computers & Electrical Engineering*, Vol. 46, pp. 328-337, Aug. 2015.
- [6] FLIR Lepton with Radiometry Quickstart Guide Document Number : 102-PS245-76 Rev 100.
- [7] Michael Beyeler, *Machine Learning for OpenCV*, Packt Publishing, 2017.
- [8] Bradski, Gary and Kachler, Adrian, "Learning OpenCV: Computer vision with the OpenCV library", O'Reilly Media, Inc., 2008.
- [9] 임동훈, "OpenCV를 이용한 영상처리", 자유아카데미, 2014.

### 저자소개

채 준 기 (Jun-gy Chae)



2019년 2월 : 한국산업기술대학교  
메카트로닉스공학과(공학사)  
관심분야 : 프로그래밍

장 지 우 (Ji-woo Jang)



2019년 2월 : 한국산업기술대학교  
메카트로닉스공학과(공학사)  
관심분야 : 설계, 프로그래밍

김 동 완 (Dong-wan Kim)



2019년 2월 : 한국산업기술대학교  
메카트로닉스공학과(공학사)  
관심분야 : 머신비전, 머신러닝

정 수 진 (Su-jin Jung)



2019년 2월 : 한국산업기술대학교  
메카트로닉스공학과(공학사)  
관심분야 : 설계, 프로그래밍

이 익 현 (Ik Hyun Lee)



2004년 2월 : 고려대학교 제어계측  
공학과(공학사)  
2008년 2월 : 광주과학기술원  
정보기전 공학부(공학석사)  
2013년 2월 : 광주과학기술원  
기전공학부(공학박사)  
관심분야 : 영상처리, 컴퓨터 비전,

정합, 융합