

COLOR RECOGNITION WEARABLE DEVICE USING MACHINE LEARNING FOR VISUALLY IMPAIRED PERSON

TAREK MOHAMED BOLAD , NIK NUR WAHIDAH NIK HASHIM*
AND NOOR HAZRIN HANY MOHAMAD HANIF

*Department of Mechatronics Engineering,
Kulliyah of Engineering, International Islamic University Malaysia,
PO Box 10, 50728 Kuala Lumpur, Malaysia.*

**Corresponding author: nikhurwahidah@iium.edu.my*

(Received: 6th June 2018; Accepted: 13th Oct 2018; Published on-line: 1st Dec 2018)

<https://doi.org/10.31436/iiumej.v19.i2.945>

ABSTRACT: Recognizing colors is a concerning problem for the visually impaired person. The aim of this paper is to convert colors to sound and vibration in order to allow fully/partially blind people to have a ‘feeling’ or better understanding of the different colors around them. The idea is to develop a device that can produce vibration for colors. The user can also hear the name of the color along with ‘feeling’ the vibration. Two algorithms were used to distinguish between colors; RGB to HSV color conversion in comparison with neural network and decision tree based machine learning algorithms. Raspberry Pi 3 with Open Source Computer Vision (OpenCV) software handles the image processing. The results for RGB to HSV color conversion algorithm were performed with 3 different colors (red, blue, and green). In addition, neural network and decision tree algorithms were trained and tested with eight colors (red, green, blue, orange, yellow, purple, white, and black) for the conversion to sound and vibration. Neural network and decision tree algorithms achieved higher accuracy and efficiency for the majority of tested colors as compared to the RGB to HSV.

ABSTRAK: Membezakan antara warna adalah masalah yang merunsingkan terutamanya kepada mereka yang buta, separa buta atau buta warna. Tujuan kertas penyelidikan ini adalah untuk membentangkan kaedah menukar warna kepada bunyi dan getaran bagi membolehkan individu yang buta, separa buta atau buta warna untuk mendapat ‘perasaan’ atau pemahaman yang lebih baik tentang warna-warna yang berbeza disekeliling mereka. Idea yang dicadangkan adalah dengan membuat sebuah alat yang dapat menghasilkan getaran bagi setiap warna yang berbeza. Disamping itu, pengguna juga dapat mendengar nama warna tersebut. Algoritma yang digunakan untuk membezakan antara warna adalah penukaran warna RGB kepada HSV yang dibandingkan dengan rangkaian neural dan algoritma pembelajaran mesin berasaskan pokok keputusan. Raspberry Pi 3 bersaiz kad kredit dengan perisian Open Source Computer Vision (OpenCV) mengendalikan pemprosesan imej. Hasil algoritma penukaran warna RGB kepada HSV telah dilakukan dengan tiga warna yang berbeza (merah, biru, dan hijau). Tambahan pula, hasil rangkaian neural dan algoritma berasaskan pokok keputusan telah dilakukan dengan lapan warna (merah, hijau, biru, oren, kuning, ungu, putih, dan hitam) dengan penukaran warna tersebut kepada bunyi dan getaran. Selain itu, hasil rangkaian neural dan algoritma berasaskan pokok keputusan mencapai hasil dapatan yang baik dengan ketepatan dan kecekapan yang tinggi bagi kebanyakan warna yang diuji berbanding RGB kepada HSV.

KEYWORDS: *colors; neural network; image processing; vibration; sound.*

1. INTRODUCTION

People with color blindness do not have the ability to distinguish between certain colors. In life, color blindness has variant negative impacts on the growth of individual patients. There are two types of color blindness, namely partial color blindness, in which the person could not distinguish certain colors, and full color blindness, where only white, grey, and black could be seen. According to the statistics, about 1 in 12 men (8%) and 1 in 200 women (0.5%) suffer from color blindness in the world. The most common form is red-green color blindness which affects men more than women [1].

Distinguishing between colors has been a serious problem faced by blind and color blind people. Problems can emerge in performing daily routine activities such as choosing and preparing food, gardening, driving a car, and selecting clothes. In addition, they are often limited in their ability to interact socially and have difficulty finding suitable employment as well. Converting colors to sound and vibration is the proposed method to combat this problem. When we look at certain colors, they can impact our emotions. For example, the color red is usually associated with aggression, orange is seen as energetic, yellow is associated with happiness, and green with nature or natural elements. The proposed system will allow a user to feel different vibrations to represent these emotions and hear the name of the color through speakers/earphones.

2. PREVIOUS WORK

There is an extensive literature on color recognition, but only a few works on converting colors to sound or vibration. Rashid et al. [2] proposed a wearable device that is able to identify colors and convert them to sound by speaking out the name of the color in two languages via wire or Bluetooth headphone. The device was developed using an Arduino Nano microcontroller. As reported by Cavaco et al. [3], color and light information could be converted into sound using a special software tool, SonarX that works in real time. This tool converts digital images into sound by mapping the HSV parameters of these image pixels into audio parameters. Palsokar [4] suggested a system that provides color perception based on RGB-HSL of the captured color image through auditory substitution. RGB color in the audio signal color model is mapped to HSL.

Manaf and Sari [5] designed and implemented a color-blind aid system. They developed finger interaction between the user and colored objects with a sound augmented reality concept. This work used two platforms: a Windows Embedded Standard 2009 (WES2009) and a system based on Windows Phone 7. Rini and Thilagavathi [6] proposed a system that can recognize cloth patterns and colors by using a support vector machine (SVM) algorithm. Identifying the color can be done using a normalized histogram of each image in the HSI color model. The features that were obtained from the image were extracted using three descriptors: radon signature and wavelet sub-bands. There were training and testing for a dataset of 627 images of four different clothing pattern designs. Trifanica, Butean, Moldoveanu and Butean [7] demonstrated that colors could be perceived into vibration with the use of an Xbox gamepad.

3. METHODOLOGY

3.1 RGB to HSV Color Conversion Algorithm

RGB and HSV are types of color space that describe how colors can be represented as tuples of numbers, normally as three values. For instance, RGB can be considered as the X, Y, and Z axes. RGB is a combination of red, green, and blue added together in many ways to reproduce array of colors that range from 0 to 255.

HSV is cylindrical-coordinate representations of points in the RGB color space. It stands for hue, saturation and value. The difference between them is that hue represents color, which has an angle from 0 degrees to 360 degrees, saturation indicates the range of grey in the color space with range from 0 to 100% and value is the brightness of the color and varies with color saturation with a range from 0 to 100% too.

The conversion was done using an image processing technique with Python™ programming language and OpenCV library. Three colors were chosen in this algorithm: red, blue, and green. The first stage was identifying the range of each color. The RGB of red color is (255, 0, 0), blue color (0, 0, 255), and green color (0, 255, 0). The second stage was taking the maximum, M , of the three RGB values of each color, which are defined as:

$$Mr = \max(R, G, B) \quad (1)$$

$$Mb = \max(R, G, B) \quad (2)$$

$$Mg = \max(R, G, B) \quad (3)$$

The third stage was taking the minimum, m of the three RGB values of each color as follows:

$$mr = \min(R, G, B) \quad (4)$$

$$mb = \min(R, G, B) \quad (5)$$

$$mg = \min(R, G, B) \quad (6)$$

Finally, calculating the HSV values:

$$V = M/255 \quad (7)$$

$$S = 1 - m/M \quad \text{if } M > 0 \quad (8)$$

$$S = 0 \quad \text{if } M = 0 \quad (9)$$

$$H = \text{Cos}^{-1} \left\{ \frac{\frac{1}{2}[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right\} \quad (10)$$

where V , S , and H are value, saturation, and hue respectively.

3.2 Artificial Neural Network Based Machine Learning

An artificial neural network is an information processing model that is inspired by the way biological neural networks function in the human brain. In this paper, multilayer perceptron (MLP) is used to train the dataset. MLP, also called feedforward neural network, consisted of more than one perceptron. They are composed of input layers (the dataset) that contain more than 1700 frames to receive the signal, an output layer that makes a decision or prediction about the inputs, and two hidden layers are between those two layers. MLP trains the input-output pairs and learns to model the correlation between those inputs and outputs. MLP is trained on the dataset that contains eight colors: red, blue, green, yellow, orange, purple, white, and black, depending on their RGB values. The purpose of neural network training is to minimize the output errors on a particular set of training data by adjusting the network weights, w . In this algorithm, OpenCV, Tensorflow and Scikit-learn libraries are used for training. Figure 1 shows the neural network diagram.

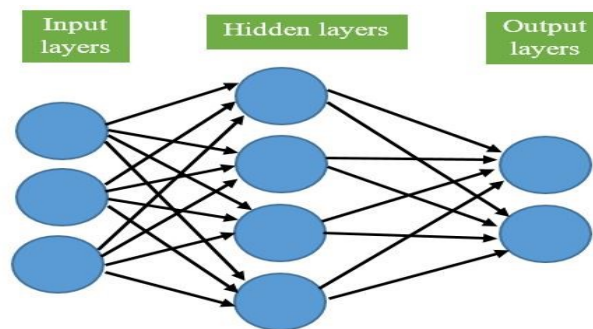


Fig. 1: Neural network diagram.

3.3 Data Preparation

Data preparation is essential for any training to ensure the dataset gives high performance and accuracy. Reshaping the images was the first step in the preprocessing stage which adjusts the images height and width. Each image contains three layers where each layer has the height of 64 pixels and width of 36 pixels. The total pixels for an image is 2304 pixels for one layer and 6912 pixels for the three layers of the image. The final step was to organize each image depending on the number of channels. Typically there are three channels of data corresponding to the colors Red, Green, and Blue (RGB). The pixel levels are usually [0,255].

3.4 Decision Tree Algorithm Based Machine Learning

The function of the decision tree is to create a training model that predicts the class or value of the target variables by learning decision rules inferred from the training dataset. The decision tree algorithm can be used for solving regression and classification problems. Classification is used to classify RGB values for each frame of the dataset for the eight colors.

3.5 Speech Synthesizer

Pyttax (Python™ text-to-speech x-platform) library in OpenCV is used to convert text to speech synthesizers. It uses different speech engines based on operating system. In this paper, 'espeak' engine is used to generate the speech, as it supports many languages. English language has been used in this paper to hear the color's name.

3.6 Vibration

The vibration actuator used in this system is a Linear Resonant Actuator (LRA) motor. This is the same vibration actuator that is used inside mobile phones. This LRA is coupled with an ultrasonic sensor and will vibrate when it detects an object within a 5 cm distance and identifies its color.

4. SYSTEM DESIGN

In this paper, five main components are used in order to run the system. The components are a Raspberry Pi3 single board computer, a Raspberry Pi camera module with 5MP resolution, an Ultrasonic sensor, an LRA vibration motor, and earphones. The design of this system was divided into three parts. The first part is a hand index box that contains the Raspberry Pi camera, Ultrasonic sensor, and a push button. The second box is attached to a stretching strap to which the vibration actuator is attached. The third part is the hand wrist box that includes the Raspberry Pi3 and a small breadboard. These parts are designed to be worn on the right hand. The prototype is shown in Fig. 2.

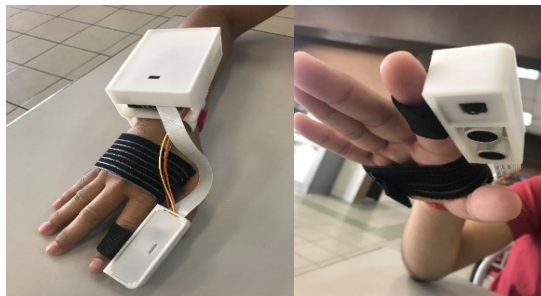


Fig. 2: Final prototype.

Figure 3 illustrates the block diagram describing the general system design. The ultrasonic sensor and the camera are the inputs of the system. The outputs are sound and vibration. The relationship between the inputs and the outputs is the image processing, which was done using the Raspberry Pi3.

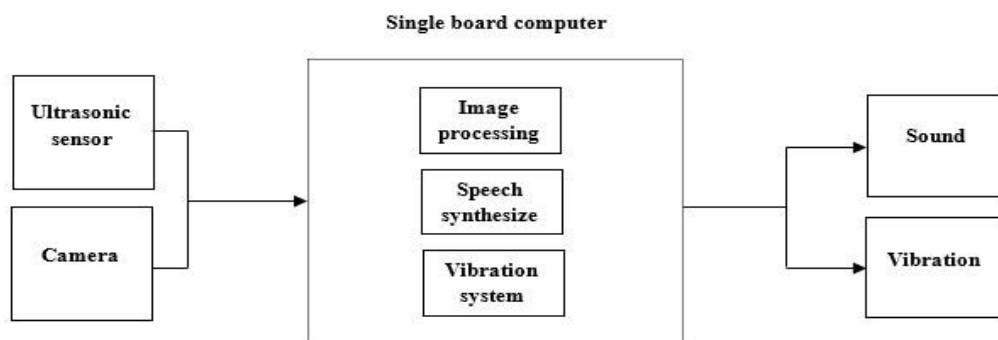


Fig. 3: System block diagram.

5. RESULTS

5.1 RBG to HSV Color Detection

A webcam camera was used in the beginning to detect just three colors: red, green, and blue. The image is converted to the RGB color space and then the equations of

RGB to HSV conversion are applied to the image to identify the color. The experiment was done by using three different objects which are blue note, red bag, and green jacket. The system was able to identify the color of each object. The detection of the camera was from a distance of approximately 2 to 4 cm for each object. Figure 4 shows the result of blue object with the HSV components.

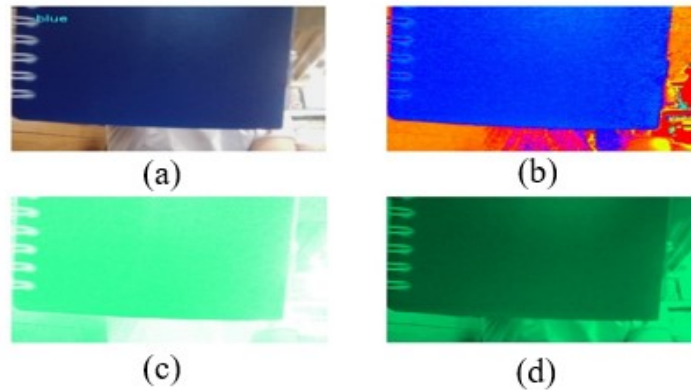


Fig. 4: Example of blue object detection using RGB to HSV. (a) HSV image, (b) Hue, (c) Saturation and (d) Value.

Fifteen different objects with different materials and colors were tested at different conditions. Each color was represented by five objects. The first group of five objects was tested in the indoor condition, the second group of five objects was tested outdoors, and the last group was tested beside an open window. Distance was measured from 2 to 3 cm for each object. The result of color accuracy is shown in Table 1.

Table 1: Result of color accuracy using RGB to HSV

Color (% accuracy)	Conditions		
	Indoor	Outdoor	Beside open window
Red	20	60	60
Green	100	60	60
Blue	100	100	100

The intensity of light affected the detection of colored objects, especially the red colored objects. The accuracies of the five red objects for indoor, outdoor and window with distance of 2 cm were the lowest among the other colors, which were 20%, 60%, and 60% respectively. Green and blue objects obtained high results as compared to the red color. It can be seen from the table that green produced the highest result within the indoor area but for outdoor area and area besides the open window produced 60% correct detection. Ultimately, blue had the best results for all conditions with an accuracy of 100%.

5.2 Machine Learning Color Detection

After using the neural network and decision tree algorithms to train the dataset, which included more than 1700 frames, the training was done by clustering each region of each frame based on RGB values for the eight colors. The experiment was conducted using a Raspberry Pi camera to classify the colors of 40 different objects in

three different conditions as well. Each color was tested on five different objects. Figure 5 shows the accuracy of each color in three different areas.

It is clear that most of the colors had good accuracy based on the neural network and decision tree algorithms. White and purple colors had low accuracy, particularly the white color, which was affected by the light. The accuracy was calculated as each object has 20% if it is the true color and 0% if it is the wrong color. Inside and outside the room were the best areas for color recognition. However, for areas inside the room, beside the window and outside the room, the accuracy of the white color are 20%, 0%, and 0% respectively, which is considered the lowest accuracy among other colors. On the other hand, black color obtained 100% accuracy for all three areas, while green obtained 100% detection for area inside the room and 80% correct detection for the remaining two areas. To sum up, the objects with black and green colors are considered the best objects with the highest accuracy as compared to white and purple colors.

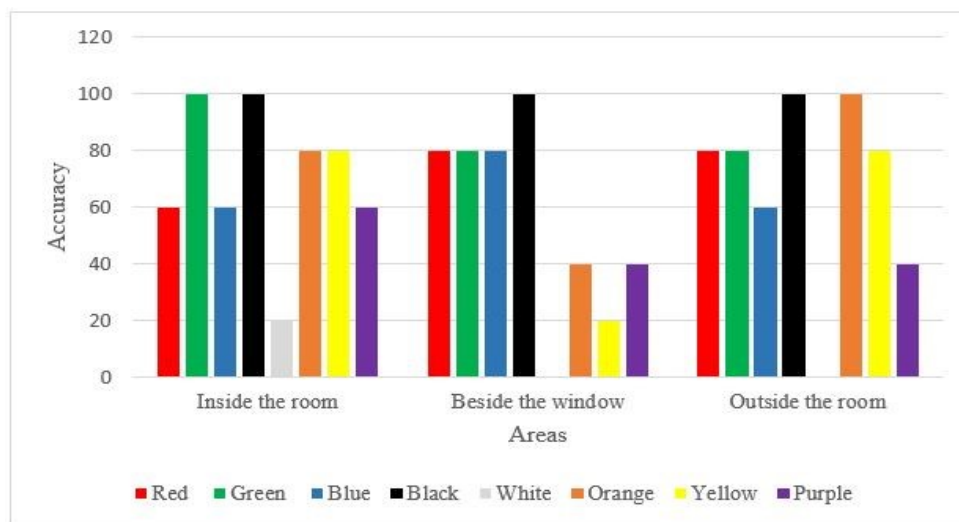


Fig. 5: Histogram bin of color accuracy using machine learning algorithm.

It is clear that most of the colors had good accuracy based on the neural network and decision tree algorithms. White and purple colors had low accuracy, particularly the white color, which was affected by the light. The accuracy was calculated as each object has 20% if it is the true color and 0% if it is the wrong color. Inside and outside the room were the best areas for color recognition. However, for areas inside the room, beside the window and outside the room, the accuracy of the white color are 20%, 0%, and 0% respectively, which is considered the lowest accuracy among other colors. On the other hand, black color obtained 100% accuracy for all three areas, while green obtained 100% detection for area inside the room and 80% correct detection for the remaining two areas. To sum up, the objects with black and green colors are considered the best objects with the highest accuracy as compared to white and purple colors.

5.3 System Setup and Integration

The last part of the experiment was connecting the RPi camera, earphones, ultrasonic sensor, and the vibration actuator to the RPi3 as follows. Firstly, the camera was turned on. Secondly, the ultrasonic, vibration, and the sound system were toggled on by the switch. From a user-experience, it can be very helpful to give additional feedback, such as a sound or vibration, to notify the user. If the measured distance is less than 5 cm, the PWM signal was set to 30 (for all 8 colors). However, if the distance is out of the range, the PWM signal

was set to zero. The result of the sound and vibration were accurate based on the colors that were detected by the camera. The speech sound uses English language. Figure 6 shows the steps and the overall system diagram.

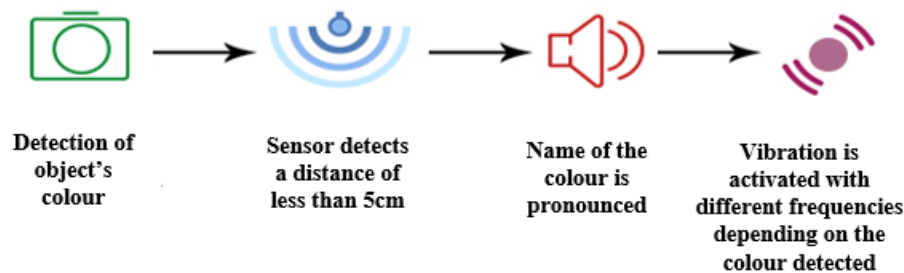


Fig. 6: System setup.

6. CONCLUSION

Hearing colors and ‘feeling’ emotions is the best way to permit color blind and blind people to distinguish between colors. Nonetheless, the most important problem that was faced in this project is the lighting surrounding the objects that significantly affected the camera during the detection of the colors. The results shown above are divided into two sections: RGB to HSV color conversion algorithm and neural network with decision tree algorithms. The neural network and decision tree algorithms have shown good performance and high accuracy as compared with RGB to HSV color conversion in recognizing colors. In addition, the sound system and vibrations were verified on the system and the results were shown to be reliable and efficient.

ACKNOWLEDGEMENT

This work is supported by the IIUM Research Initiative Grant Scheme (RIGS16-071-0235).

REFERENCES

- [1] Ohkubo T, Kobayashi K. (2008). A color compensation vision system for color-blind people. Proceedings of the SICE Annual Conference, 1286–1289. <https://doi.org/10.1109/SICE.2008.4654855>
- [2] Rashid H, Al-mamun ASMR, Sijanur M, Robin R, Ahasan M, Reza SMT. (2016). Bilingual Wearable Assistive Technology for Visually Impaired Persons. <https://doi.org/10.1109/MEDITEC.2016.7835386>
- [3] Cavaco S, Mengucci M, Henriques T, Correia N, Medeiros F. (2013). From pixels to pitches : unveiling the world of color for the blind. <https://doi.org/10.1109/SeGAH.2013.6665305>
- [4] Palsokar A. (2016). A RGB-HSI Conversion Based Model to Map the Colour Perception Into Audio Equivalent to Assist Visually Impaired Persons. <https://doi.org/10.1109/IACC.2016.96>
- [5] Manaf A, Sari R. (2011). Color Recognition System with Augmented Reality Concept and Finger Interaction. <https://doi.org/10.1109/ICTKE.2012.6152389>
- [6] Rini JJ, Thilagavati B. (2015). Recognizing clothes patterns and colours for blind people using neural network, 1–5. <https://doi.org/10.1109/ICIIECS.2015.7193006>
- [7] Trifanica A, Butean A, Moldoveanu A, Butean D. (2015). Gamepad Vibration Methods to Help Blind People Perceive Colors, 37–40. Retrieved from <https://www.researchgate.net/publication/283070958>