Original Article

e-ISSN: 2581-0545 - https://journal.itera.ac.id/index.php/jsat/

Received 10th November 2020
Accepted 3rd December 2020
Published 20th December 2020

Open Access

DOI: 10.35472/jsat.v4i2.344

The Application of *Multilevel Colour Thresholding* for Coffee Colour Detection in a Coffee Maker Prototype

Nova Resfita *[a], Rahmadi Kurnia [b], Fitrilina [c]

[a] *Department of Biomedical Engineering, Institut Teknologi Sumatera*

[b] *Department of Electrical Engineering, Universitas Andalas*

[c] *Department of Electrical Engineering, Universitas Bengkulu*

** Corresponding E-mail:* nova.resfita@bm.itera.ac.id

Abstract: The development of computer vision has expanded widely as there is a vast number of its applications in various aspects of daily life. One of its implementations is integrating the image processing technique on a prototype coffee machine based on the speech recognition system. This study aims to detect the requested coffee colour spoken by users which are *black, middle* and *light*. The sensor used in this research is a digital PC camera and the applied method is *Multilevel Colour Thresholding*. Of all experiments conducted, the image processing technique can work perfectly as the camera is able to identify the requested colour of the coffee solution. Furthermore, the system might be developed by improving the multilevel colour thresholding technique as well as advancing the hardware design in order to establish more robust coffee machine based on the requested colour.

Keywords: *image processing, multilevel colour thesholding, video segmentation*

Introduction

Human-computer interaction has widely developed over decades in which it has improved various technologies to ease the people activities. One of the most leading advancement in daily technologies is the implementation of image processing to particular machines.

In principle, computer vision aims to imitate the human visual system to perform particular tasks by incorporating optical device such as camera to capture an image or record the video [1]. Computer vision consists of two main parts, the image processing and pattern recognition [2]. In visual machines, the recording data cannot be directly interpreted by the machine. Hence, the image processing step is the key point of interpreting the given data before a task is executed [3].

The application of multilevel colour thresholding technique has widely used in various research [4]. Jan Kubicek et. al applied similar method to extract the blood vessels [5], Kumar et.al identified the defective region in leather surfaces [6], and Harrabi et. al detected the breast cancer images [7]. Additionally, based on the previous research conducted by Aksara [8], an object can be tracked by applying the static webcam according to its colour information. Another similar research was also conducted by Fitrilina et. al to examine the mobile robot based on the colour detection to be commanded by speech recognition system [9]. Furthermore, due to wide preferences in making coffee, it is benevolent to utilise the camera on the coffee machine. Thus, this research aims to implement the use of an image processing technique to control the coffee machine based on the colour spoken in order to serve a cup of coffee.

The whole system consists of an integrated speech recognition module, image processing module and microcontroller device together with the application of DC and stepper motor. The coffee machine works according to the spoken colours, which are “*black*”, “*middle*” and “*light*”. These words then activate the motors controlled by a microcontroller. When the camera detects the desired spoken colour of the coffee, the microcontroller will turn off the motors, so then the coffee is ready. In addition, this paper will focus on the colour detection by the image processing module.

Method

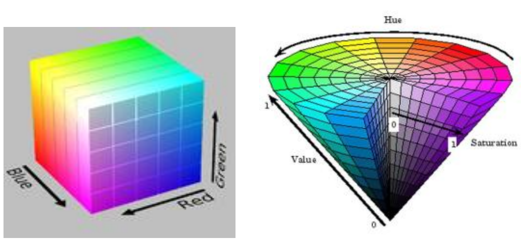
Image and Video Processing

Image is a two-dimensional representation, similarity or imitation of an object that contains particular information [10]. Mathematically, it is a continuous function of light intensity on two-dimensional planes (x,y) [11]. Furthermore, video is a series of still images (frames) displayed sequentially at a certain time interval with particular speed to give the impression of a moving image to the eyes. It consists of arranged pixels on a three-dimensional array in which the two dimensions represent the spatial images vertically and horizontally and one dimension indicates the time or number of images. To analyse a video, it is necessary to divide the video into smaller units that have semantic relations such as shot, key frame and scene [12].

Video Segmentation

Segmentation is a separation process of an image based on its specific areas holding particular homogeneity such as texture, colour and intensity. Its purpose is to differentiate the object and background of an image, so it would be convenient to detect the desired object [12]–[16]. In this research, the videos recorded were segmented based on its frame detection and value range of its colour model representation.

The video formed by series of images consists of a particular range of colour spectral information [10]. This range is depicted on the colour models, such as RGB (Red, Green, and Blue) and HSV/HSI (Hue, Saturation, and Value/Intensity). Figure 1 depicts the discrepancy between RGB and HSV colour systems.



RGB colour Model HSV colour Model

Figure 1. Comparison between RGB and HSV Model [17], [18]

As the colour components of an object in RGB colour space are hardly separated due to the correlations between R, G and B component, it is highly essential to convert them to HSV colour space. Furthermore, the chrominance components (hue and saturation) mimick the human perception of colour which facilitates the image processing better [18]. In addition, Vandenbroucke [19] proposed one of the colour space classifications to quantify subjective human colour perceptions by converting the image from RGB to HSV by:

$$H = \begin{cases} 0, \left(60^\circ \times \dfrac{G-B}{Max-Min} + 360^\circ\right), & if\ Max = Min \\ \times\ mod\ 360^\circ, & if\ Max = R \\ \left(60^\circ \times \dfrac{B-R}{Max-Min} + 120^\circ\right), & if\ Max = G \\ \left(60^\circ \times \dfrac{R-G}{Max-Min} + 240^\circ\right), & if\ Max = B \end{cases}$$

$$S = \begin{cases} 0, & if\ max = 0 \\ \dfrac{Max-Min}{Max} & otherwise \end{cases}$$

$$V = Max \tag{1}$$

Multilevel Colour Thresholding

Multilevel colour thresholding is a segmentation technique applying a wide range of colour spaces [7], [8], [20]. In this study, the colour spaces applied are RGB converted to HIS as the most important parameter is hue. During the thresholding process, the system will pass the hue colour of particular parts of the image according to its determined range. These range then are segmented as an object and the rest is the background. In order to detect the desired coffee colour by the camera, Table 1 lists the ranges of colour components of the coffee:

Table 1. Colour Components Range of Coffee in HIS Colour Space.

No	Type of Coffee	Colour	HIS Range (Hue, Saturation, Intensity)
1	Black		1-25; 0-30; 10-240
2	Middle		1-25; 31-120; 20-240
3	Light		1-25; 141-240; 10-240

Based on HSI (Hue, Saturation and Intensity) range listed in Table 1, all those three types of coffee colour have similar range of hue value as they have the same basic coffee. "*Black*" represents the coffee with no or little creamer, meanwhile "*middle*" is the coffee with slightly more creamer than "*black*", and "*light*" has the highest amount of creamer. There was no exact measure of creamer as the whole system worked automatically to pour the creamer into the glass and stirred it until the requested coffee colour is achieved.

In this research, the window size to capture the video was 320 (height) x 240 (width) pixels and the video format was *.avi*. The video recording was taken from frame to next frame, starting from the point 0,0 *(i,j)* to point *(m,n)* in which *i,j* represented the row and column respectively, and *m,n* depicted the frame height and width. The pixels on the point *(i,j)* then was defined as a true colour (recognised in 24 bits [20]), where the first 8 bits were blue, the second 8 bits were green and the last 8 bits were red. Subsequently, these RGB colours were converted into the HSI colour model. The threshold for each parameter (Hue, Saturation and Intensity) was defined as listed in Table 1. The pixels included in the specified range were detected as the object, meanwhile those outside the range were portrayed as the background.

To proceed this experiment, the hardware setup was designed in prototype machine where the camera (*Mediatech USB Digital PC camera with frame rate 320 x 240 pixels up to 30 frames/sec*) was placed horizontally to the coffee glass (transparent). Additionally, the DC motor rotated the mixer to stir the coffee solution and stepper motor moved accordingly to pour the creamer into the drink. Both DC and stepper motor were controlled by microcontroller ATMega 8535. As the camera recorded the video and detected the requested colour, the whole system would stop and the coffee was served.

Results And Discussion

The experiments started by executing the speech processing module in which the requested type of coffee drink was spoken and processed through the speech recognition system. The recognised word (either "*black", "middle" or "light"*) then became the input to the image processing and microcontroller systems. Based on this word, the hardware module worked accordingly by activating the microcontroller to control the DC motor stirring the coffee mixer and stepper motor to manage the cream pouring into the coffee solution. At the same time, the camera was also activated to record the video of the coffee colour. When the camera identified the requested colour, the microcontroller would terminate the DC and stepper motor activity, then the coffee is ready.

The experiments were conducted according to the trials of speech recognition system. Of 45 tests (15 attempts for each coffee colour) of a trained speech sources, the success rate was 95.96%, which 2 trials for the *"light"* coffee was wrongly executed. In additon, this success rate was declined to 80,67% (of 150 trials) by the untrained sources in which 2 attempts were false for *"black"*, 4 trials in *"middle"* and 27 mistakes in *"light"*.

Based on those executed trials, results of segmentation by applying multilevel colour thresholding on this research are displayed in Figure 2:

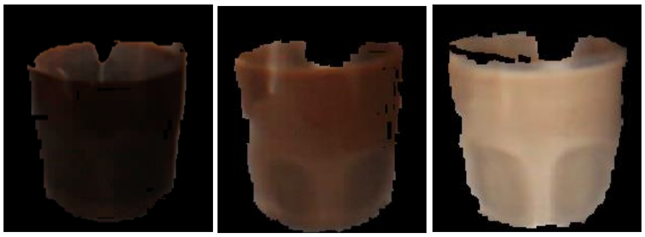


Figure 2. Segmentations of the coffee colours (*black, middle, and light*)

Figure 2 shows that the system is capable to separate the object (coffee colours) from its background by selecting the colours determined on HIS range and blackening the background based on the threshold value set by the system for each object. However, the coffee colour detection is highly dependent to the recognised spoken word. For instance, if a person speaks *"black"* and the speech recognition system identifies it as *"middle"*, then the camera will consider *"middle"* as the input and detect the colour in range of *"middle"*.

Despite the image processing module can work accordingly, the segmentation process is not perfectly performed as the presence of noise cannot be

Published by: Lembaga Penelitian, Pengabdian Masyarakat, dan Penjaminan Mutu Institut Teknologi Sumatera, Lampung Selatan, Indonesia

segmented as well. It happens due to the influence of light on the object creating the appearance of a specific colour of the background that is similar to the targeted object or the existence of other objects having an identical colour that can be detected by the camera system.

In general, the image processing system to detect the coffee colour can impeccably perform its tasks as the success rate is 100%. The system is able to identify the coffee solution based on the recognised speech, either it is "*black*", "*middle*" or "*light*". Nevertheless, due to the ability of speech recognition systems, some requested coffee might be wrongly executed by the system.

The implementation of multilevel colour thresholding in a prototype coffee machine has achieved its purpose in detecting the coffee colour. The system was sufficiently reliable to distinguish the coffee solution with its background, although some drawbacks appeared on the system. The enhancement of Multilevel Colour Thresholding method might be proceeded by applying advanced filtering and machine learning. In addition, the camera can also be possibly mounted vertically so the colour coffee might be detected directly without the media (transparent glass) as it has some light effects. Furthermore, the background should also be set in uniform colour with a preference, either black or white. Thus, the image processing to detect coffee colour can be well executed.

Conclusions

The aim of this research to detect coffee colour by applying Multilevel Colour Thresholding with a camera sensor on a coffee maker prototype based on speech recognition system can be completely accomplished as the success rate of detection is 100% even though the rate of the recognised words by speech recognition system were 95.96% and 80.67% for the trained and untrained sources respectively. Furthermore, in order to improve the segmentation process, a filtering is required to remove noise and eliminate other objects that are relatively smaller than the desired target. The system can also be developed by enhancing the image processing technique and improving the hardware design.

Conflicts of interest

There are no conflicts to declare

References

[1] T. S. Huang, "Computer Vision: Evolution and Promise," *Report*, 1997.

[2] O. Cosido *et al.*, "Hybridization of convergent photogrammetry, computer vision, and artificial intelligence for digital documentation of cultural heritage-A case study: The magdalena palace," *Proc. - 2014 Int. Conf. Cyberworlds, CW 2014*, no. August, pp. 369–376, 2014, doi: 10.1109/CW.2014.58.

[3] V. Wiley and T. Lucas, "Computer Vision and Image Processing: A Paper Review," *Int. J. Artif. Intell. Res.*, vol. 2, no. 1, p. 22, 2018, doi: 10.29099/ijair.v2i1.42.

[4] X. Bao, H. Jia, and C. Lang, "A Novel Hybrid Harris Hawks Optimization for Color Image Multilevel Thresholding Segmentation," *IEEE Access*, vol. 7, pp. 76529–76546, 2019, doi: 10.1109/ACCESS.2019.2921545.

[5] J. Kubíček *et al.*, "Modeling and objectification of blood vessel calcification with using of multiregional segmentation," *Vietnam J. Comput. Sci.*, vol. 5, 2018, doi: 10.1007/s40595-018-0122-z.

[6] M. Praveen Kumar and S. Denis Ashok, "A multi-level colour thresholding based segmentation approach for improved identification of the defective region in leather surfaces," *Eng. J.*, vol. 24, no. 2, pp. 101–108, 2020, doi: 10.4186/ej.2020.24.2.101.

[7] R. Harrabi and E. Ben Braiek, "Color image segmentation using multi-level thresholding approach and data fusion techniques: Application in the breast cancer cells images," *Eurasip J. Image Video Process.*, vol. 2012, no. 1, pp. 1–11, 2012, doi: 10.1186/1687-5281-2012-11.

[8] A. C. Yuda, "Object Tracking Pada Gerakan Non-Linier Berdasarkan Informasi Warna," Universitas Andalas, Padang.

[9] . F., R. Kurnia, and S. Aulia, "Pengenalan Ucapan Metoda MFCC-HMM Untuk Perintah Gerak Robot Mobil Penjejak Identifikasi Warna," *J. Nas. Tek. Elektro*, vol. 2, no. 1, pp. 31–40, 2013, doi: 10.20449/jnte.v2i1.95.

[10] Y. Wang, J. Ostermann, and Y.-Q. Zhang, *Video processing and communications*. Prentice Hall, 2002.

[11] A. McAndrew, *A computational introduction to digital image processing, second edition*. 2015.

[12] Y.-J. Zhang, "An Overview of Image and Video Segmentation in the Last 40 Years," *Adv. Image Video Segmentation*, pp. 1–16, 2011, doi: 10.4018/978-1-59140-753-9.ch001.

[13] D. Khattab, H. M. Ebied, A. S. Hussein, and M. F. Tolba, "Color image segmentation based on different color space models using automatic GrabCut," *Sci. World J.*, vol. 2014, 2014, doi: 10.1155/2014/126025.

[14] S. Pare, A. Bhandari, A. Kumar, and G. K. Singh, "An optimal Color Image Multilevel Thresholding Technique using Grey-Level Co-occurrence Matrix," *Expert Syst. Appl.*, vol. 87, 2017, doi: 10.1016/j.eswa.2017.06.021.

[15] S. Sarkar, S. Das, and S. S. Chaudhuri, "A multilevel color image thresholding scheme based on minimum cross entropy and differential evolution," *Pattern Recognit. Lett.*, vol. 54, pp. 27–35, 2015, doi: 10.1016/j.patrec.2014.11.009.

[16] C. Y. Tsai and T. Y. Liu, "Real-time automatic multilevel color video thresholding using a novel class-variance criterion," *Mach. Vis. Appl.*, vol. 26, no. 2–3, pp. 233–249, 2015, doi: 10.1007/s00138-014-0655-9.

[17] W. Syahrir, A. Suryanti, and C. Connsynn, "Color grading in Tomato Maturity Estimator using image processing technique," *2009 2nd IEEE Int. Conf. Comput. Sci. Inf. Technol.*, pp. 276–280, 2009.

[18] N. Ibraheem, R. Z. Khan, and M. Hasan, "Comparative Study of Skin Color based Segmentation Techniques," *Int. J. Appl. Inf. Syst.*, vol. 5, pp. 24–39, 2013, doi: 10.5120/ijais13-450985.

[19] N. Vandenbroucke, L. Macaire, and J.-G. Postaire, "Color image segmentation by pixel classification in an adapted hybrid color space. Application to soccer image analysis," *Comput. Vis. Image Underst.*, vol. 90, no. 2, pp. 190–216, 2003, doi: https://doi.org/10.1016/S1077-3142(03)00025-0.

[20] V. Haghighatdoost and R. Safabakhsh, "Automatic Multilevel Color Image Thresholding by the Growing Time Adaptive Self Organizing Map," pp. 1768–1772, 2006, doi: 10.1109/ictta.2006.1684653.

Published by: Lembaga Penelitian, Pengabdian Masyarakat, dan Penjaminan Mutu Institut Teknologi Sumatera, Lampung Selatan, Indonesia