

LOW-COST DOME FOR MULTI-PURPOSE DIGITALIZATION ON SMALL OBJECT

AUFACLAV ZATU KUSUMA FRISKY^{1,*}, DIYAH UTAMI KUSUMANING PUTRI¹
MAGESH CHANDRAMOULI², DANDY ZICKY DIVALDY¹, REMAREZI RAFSANJANI¹
ANDI DHARMAWAN¹ AND AGUS HARJOKO¹

¹Department of Computer Science and Electronics
Universitas Gadjah Mada

Bulaksumur, Yogyakarta 55281, Indonesia
{ diyah.utami.k; andi_dharmawan; aharjoko }@ugm.ac.id
{ dandyzicky03; remarezirafsanjani2020 }@mail.ugm.ac.id
*Corresponding author: aufaclav@ugm.ac.id

²Department of Computer Information Technology and Graphics
Purdue University Northwest
Hammond, IN 46323, USA
magesh@pnw.edu

Received December 2022; accepted March 2023

ABSTRACT. *The production of digital visual records via computer vision is becoming more accessible. In this study, we build a low-cost multipurpose dome capable of supporting a number of technological applications, including the necessity for consistent illumination on Optical Character Recognition (OCR), Multispectral Imaging, and Photometric Stereo. To achieve the best results possible for each process, a dome is constructed to contain the environment. Nevertheless, each application procedure necessitates the production of a unique dome. The dome is designed affordably by manipulating the position, brightness, and color of Light-Emitting Diode (LED) sources. The size of a moveable dome has become a significant alternative for digitizing small objects, which may then be duplicated for educational purposes. The proposed dome is also linked to an application with a Graphical User Interface (GUI) that allows the user to modify settings and control the dome LED. The outcomes of our experiments indicate that our illumination dome can be applied on three different applications. Despite its low cost, these results demonstrate the dome's versatility and potential applications in academic research settings and industrial machine vision. In future, this dome can be used to produce a small object digital acquisition.*

Keywords: Low-cost dome, Digitalization, Photometric stereo, OCR, Multispectral imaging, Multi-purpose

1. Introduction. The main aspect of many applications of computer vision algorithms, such as photometric stereo [1, 2], multispectral imaging, and also text binarization, is how a particular scene is lit [3]. If the lighting situation is not well managed, the application is likely to produce inconsistent results [4]. The photometric stereo reconstruction of a 3D object, for instance, employs numerous 2D pictures with various light sources [1]. To display this information on numerous light sources, a light-controllable dome is required. Additionally, the lighting dome works as a barrier against ambient light noise from the exterior environment. Red, Green, and Blue (RGB) Light-Emitting Diode (LED) arrays are affixed throughout the interior of the dome to serve as changeable point light sources. After that, photometric stereo may be applied to the collected images to generate the mesh file, which is the 3D model's structure. Using this method, Brenner et al. [2] created a

mesh of Roman coin data. Their method uses 54 types of light to produce an accurate 3D model.

Alternately, the RGB LED can emit nearly all wavelengths of the visible light spectrum by varying its color and may be utilized for inexpensive multispectral imaging. Multispectral imaging is the recording of light intensity as a function of wavelength at each pixel or geographic location of an image [5]. Due to the dome's homogenous lighting, it can also improve the accuracy of Optical Character Recognition (OCR) on document pictures [6]. A suitable brightness level is one of the most significant factors in achieving the best outcome. Depending on the quality of the paper and the color of the ink used, each document requires a unique, specific brightness [7, 8].

Based on the aforementioned method, each approach requires a different configuration of the dome to be effective. The construction of multiple domes is thus both necessary and inefficient. Basically, the shape of the required dome is the same; it only needs a system that can control the LED based on requirements. Furthermore, a compact dome that can fulfill all the requirements and be controlled by the user is required. After we present brief results from previous research on the relevant topics, we conduct experiments with each of the respective topics using our device to demonstrate its capability while also explaining the methodology we use. We end with a conclusion in which we talk about the limits of our study and possible directions for future research in this area.

2. Related Works. One way to build a multi-functional illumination device that is suitable for static objects is to use a dome-shaped setup. The illumination setup by Gong et al. [9] shows that dome illumination eliminates reflection issue for a non-Lambertian surface such as glass. For 3D reconstruction via photometric stereo, a lighting system that can vary the direction of light in order to capture the required images is needed. The surface normal can be calculated using several images of a surface taken in different light directions. Hayakawa [10] demonstrated that if illumination directions are unknown, the surface normal may be reconstructed up to a general linear transformation. This generic linear ambiguity may be reduced to a 3D rotation ambiguity if one can identify six lights with similar intensities or six normals with equal albedo. From the restored surface normal, the estimated surface depth may be calculated.

The illumination dome at UCL [11] allows for the acquisition of a series of images of an object from a fixed zenithal camera position, with illumination provided by 64 flash lights at known coordinate positions on the hemisphere [12, 13, 14]. Zhou and Tan [15] created a ring light setup for photometric stereo purposes. It was shown that the surface normal of a Lambertian scene can be recovered with at least five lights on a cone-shaped structure for up to two types of rotations and a scaling compounded with a mirror ambiguity.

For multispectral imaging applications, equipment consisting of a filter-wheel camera [16] was constructed. Combining the pictures obtained via the application of a series of bandpass filters leads to the recording of the multispectral data cube. The need for constant illumination is also required for OCR [7]. This technology is used to transform image-based transcribed, handwritten, or printed text documents into text data that can be altered and reused [17, 18]. Most existing OCR techniques presume that the input image was captured in optimum lighting circumstances and only focus on text detection and recognition, so it is not rare for a system to use extra algorithm such as contrast enhancement [19] or shadow removal such as in this study [20]. From the aforementioned works, it is evident that dome instruments were typically employed for their own algorithms with specific constraints. Our illumination dome is designed to be versatile and cost-effective. Consequently, our apparatus demonstrated that, for its intended purpose, it is applicable to three different digitalization algorithms.

3. Proposed Dome Setup. In this study, we build a low-cost multipurpose dome capable of supporting a number of technological applications, including the necessity for consistent illumination in OCR [21, 22], multispectral imaging [23], and photometric stereo [24]. The dome illumination setup consists of a hollow dome construction that has LED strips mounted inside of it, as well as a camera that is suspended 50 centimeters above the dome’s center. For the electronics, a 720W power supply unit is used, 12V LED strips are organized in arrays, and an Arduino board is used to generate Pulse Width Modulation (PWM) signals. These signals determine the colors that are emitted and the brightness of the LEDs [25]. Based on the state of the digital pin, the Arduino board controls the selection of the LED (which LED to turn on). The system is controlled by a GUI connected to a Bluetooth serial module. In addition to this, we printed a circuit board to provide power to the Arduino, manage all of the Arduino pins that are connected to the LEDs, and also boost the PWM signals for a more stable output. The technical component of the proposed dome is listed in Table 1 and Table 2. The LED placement inside the dome and the schematic of the proposed dome from the side view and bottom perspective, respectively, are depicted in Figure 1.

TABLE 1. Technical components of the illumination dome

Number of possible illumination direction	10
Input voltage	220VAC
Control mechanism	Arduino PWM
Max power usage	210W
Total cost in USD (without camera)	\$149

TABLE 2. Arduino Mega parameters (ATmega2560)

Microcontroller	ATmega2560
Operating voltage	5V
Input voltage (recommended)	7-12V
Input voltage (limits)	6-20V
Digital I/O pins	54 (of which 14 provide PWM output)
Analog input pins	16
DC current per I/O pin	40mA
DC current for 3.3V Pin	50mA
Clock speed	16MHz

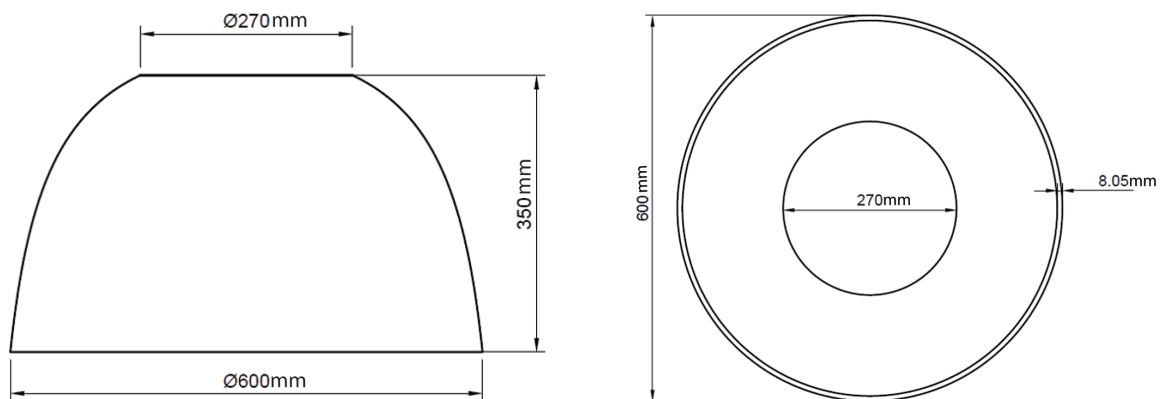


FIGURE 1. Dome design, left = side perspective, right = bottom perspective

Currently, our Printed Circuit Board (PCB) is capable of powering 20 LED arrays. For an Arduino Mega, the maximum current per pin is 40mA. Furthermore, in order to boost the PWM signal, which controls the light intensity and color to help it deal with resistance, we designed a simple PWM current amplifier using S9013 NPN transistors and 2N7000 N-Channel MOSFETs as drawn in the schematic in Figure 3.

With a diameter of 800mm and a height of 350mm, our domes are fairly portable. The LEDs installed in the dome wrap completely around the inside of the dome. Inside are 20 panels, whose location, light intensity, and color may be manually modified. A PWM controller with an intuitive GUI is created for the settings (see Figure 2).

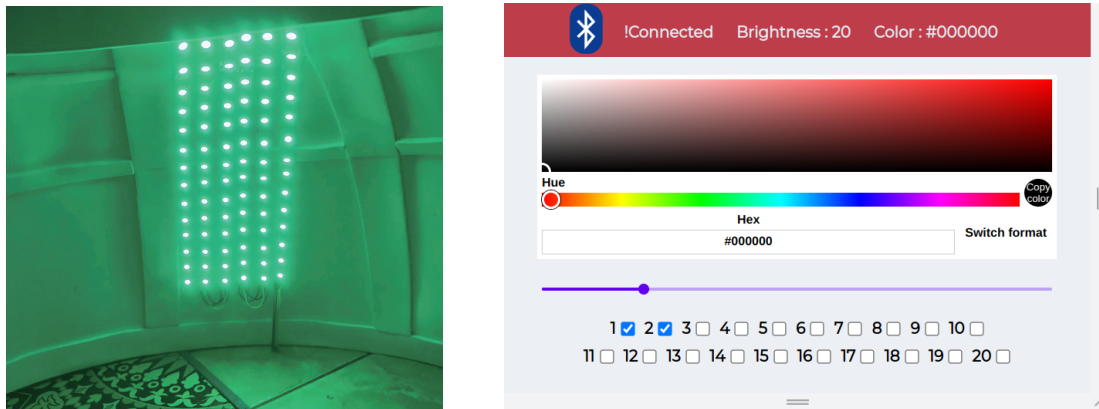


FIGURE 2. LED array placement and the Graphical User Interface to control them

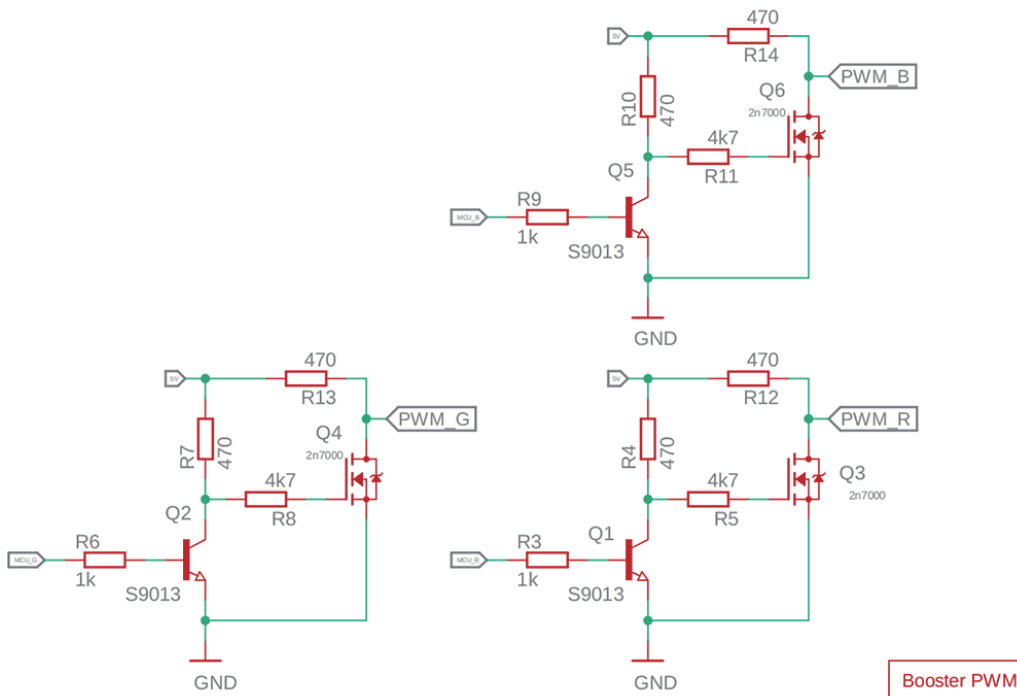


FIGURE 3. Schematic for PWM signals conditioning

For powering and selecting which LED array to activate, the IRF540 N-Channel MOSFETs are used to act as a switch for controlling LEDs' color and intensity by adjusting the PWM of each color pins. When the MOSFET is turned on, 12V voltage is applied to the R/G/B pin on the LED. For selecting the panel to be turned on, a PNP transistor, TIP32, is used to cut or connect power into the LED's panel (see Figure 4).

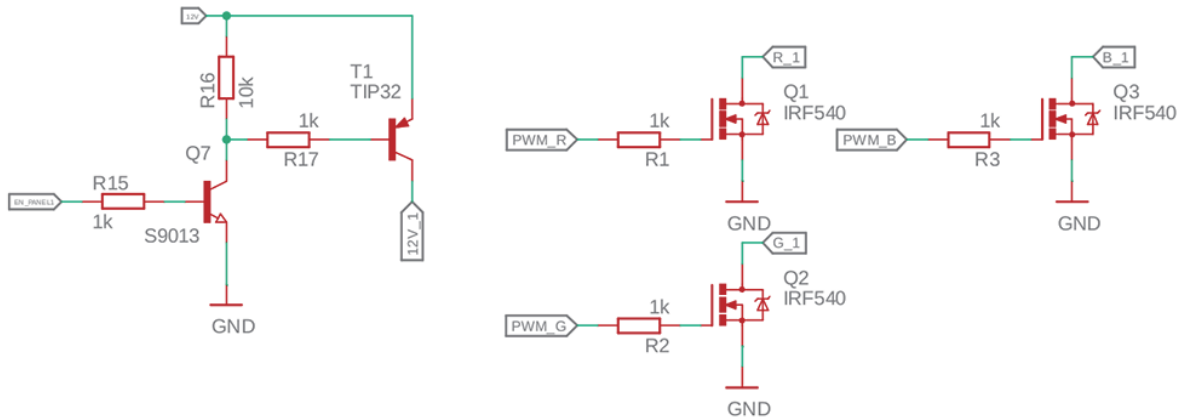


FIGURE 4. Schematic for the switching of the LEDs

To capture the image, we used a 12-megapixel camera on a smart phone for all of the experiments. The capturing procedure will not be handled by the GUI, but rather by the user on the smart phone. Because the white light generated by the LED is on the cool side, we used higher color temperature settings for the camera to compensate. We use ISO 80, shutter speed 1/125, color temperature 5550K, focal length 26mm, and aperture f1.8. The main mechanism of communication between the Arduino and the GUI is handled by the Bluetooth serial module HC-05. The module acts as a serial port, allowing the web serial API to connect with the Arduino via serial communication. There are 3 pieces of data to send, which are the light brightness, color, and an array of the states of the LED arrays in 0 and 1. A callback function implemented in JavaScript then executes whenever a change is detected for real-time effect, which sends a stream of JSON containing all three pieces of data. The Arduino receives the JSON stream and parses the serialized message into its original data format. It allows the Arduino to write the PWM pin duty cycle, which controls the specified color, intensity, and the array of enabled pins to activate a particular LED array.

4. Result and Discussion. There are three experiments that will be covered in this paper: photometric stereo, multispectral imaging, and also OCR. All the experimental data were obtained using the same dome with different configurations. Further detail about all experiments is explained in the following subsections.

4.1. Photometric stereo. This experiment utilized twenty images of the Indonesian Rp500.00 coin that were captured using the illumination dome. The main photometric stereo 3D reconstruction algorithm used is uncalibrated photometric stereo [26]. The program took approximately 50 seconds to complete surface normal calculation, depth map estimation, and 3D mesh export to a .obj file (see Figure 5).

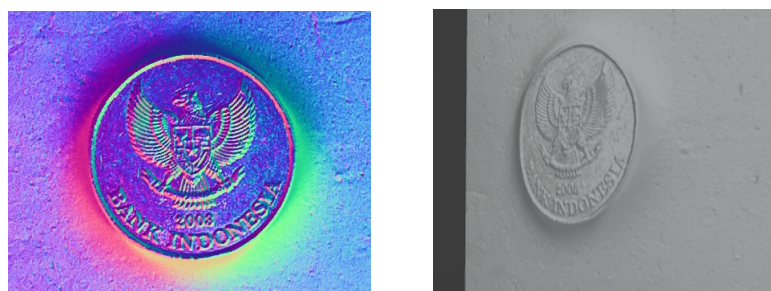


FIGURE 5. Photometric results, left = surface normal and right = 3D reconstructed model

Figure 5 illustrates the estimated surface normal map obtained through Singular Value Decomposition (SVD) [27]. For this experiment, the image size is compressed to 384×512 for faster processing. With shadows, inter-reflections, and non-Lambertian effects ignored, the image intensity matrix can be formulated as $I = NL$. Here, I is an $n \times 3$ matrix formed by pixel intensities, with $n = 1996608$ because each image contains 1996608 pixels. N and L are $n \times 3$ and 3×3 matrices, respectively. Each row of N indicates the scaled surface normal (unit surface normal multiplied by albedo), and each column of L is the scaled lighting direction (unit lighting direction multiplied by its intensity). In uncalibrated photometric stereo, only I is known, and both N and L are unknown. Applying the SVD, the matrix I can be decomposed as

$$I = UWV^T = (UW^{1/2}) (W^{1/2}V^T) \quad (1)$$

To aid in computation of U , W , V^T , the OpenCV SVD class is used. Matrix V^T is transposed back to matrix V because it contains the eigenvectors of $I^T I$, which practically we can treat per column respectively as the red, blue, and green components for the surface normal map N in 8-bit RGB channel.

As we can see in Figure 5, the non-Lambertian feature of the coin results in minimal artifacts, which is to be expected given that just three images are used, and the fundamental objective of reconstructing the coin in three dimensions is still achieved.

4.2. Multispectral imaging. Seven images of the 5 Turkish Lira banknote are captured using seven different colors. All the images are split into three single channels (R, G, and B) in grayscale, so there are a total of 21 images. Next, the images are transformed into a one-dimensional array and fitted into a matrix A with the size of $21 \times n$ with n being the image pixels. From there, we applied PCA [28] to getting the first K principal component vectors. Firstly, we need to find $C = A - a$ in which C is the matrix from the subtraction of each column of A with the mean vector a . So, the covariance matrix is

$$S = C \cdot C^T \quad (2)$$

The principal components are the eigenvectors in decreasing order of the eigenvalues [28]. The corresponding eigenvalues can be used to estimate the relative significance of the various principal component vectors. For that reason, we only take the first four principal component vectors as an example because they have the highest explained variance ratio, which is the percentage of variance that is attributed to each of the selected components (see Figure 6).



FIGURE 6. First four highest principal component vectors projected

4.3. Text recognition. The Python module pytesseract [6] was used for the OCR process. Before the text recognition, there is some preprocessing to prepare the document, which includes image masking and Otsu binarization. The Otsu approach [29] is a variance-based strategy for determining the threshold value with the smallest weighted variance between the foreground and background pixels. After that, pytesseract model is used to detect characters in binary images based on pixel distribution analysis and training. At this stage, all outlines are collected and organized as blobs. Blobs are arranged in the form of text lines, after which the areas and lines are analyzed and adjusted into a proportional text form. Based on the type of character spacing, lines of text are divided into words using definite spaces and fuzzy spaces. The recognition stage is then continued using letter shape recognition with a high degree of confidence in a stage known as adaptive recognition (first pass).

An old manual page written in German with 572 words is used. From the OCR result, Figure 7 demonstrates that only four errors resulted (from 572 words), indicating a 99.3% accuracy for the OCR process. The circular light sources from our illumination dome setup have successfully eliminated most of the shadows formed on the document, hence the near-perfect accuracy. The LEDs inside the dome are also controlled so that the right brightness for the camera is achieved, preventing overexposure or glare.

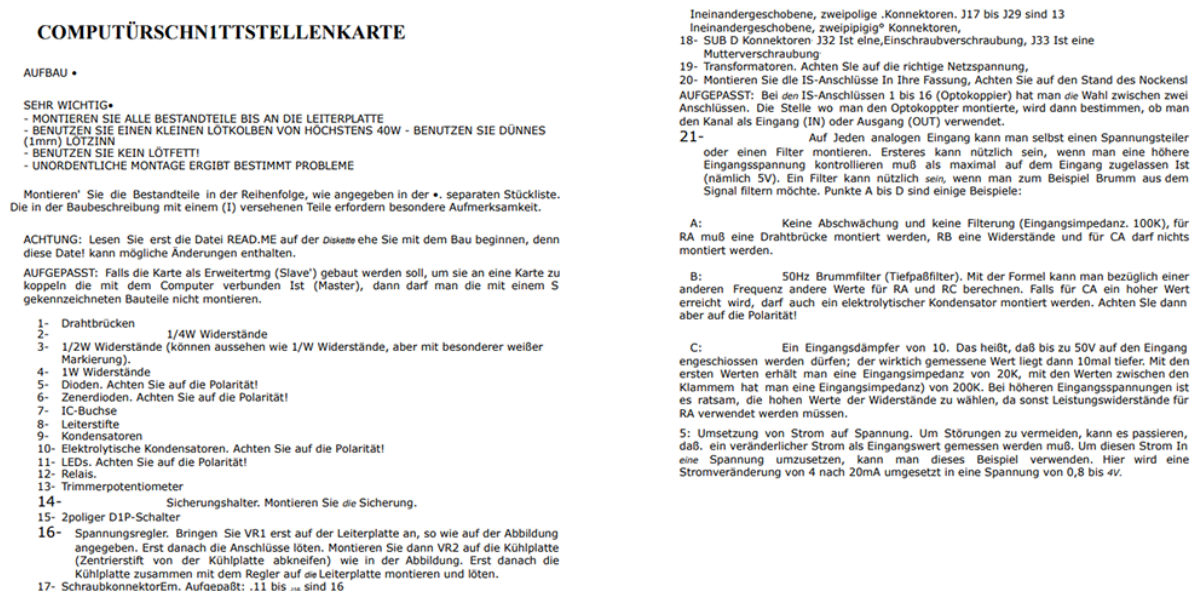


FIGURE 7. Text recognition via OCR

5. Conclusions. The outcomes of our experiments indicate that our illumination dome can be applied to three different computer vision applications. Despite its affordability, these results demonstrate the dome's versatility and potential applications in academic research settings as well as industrial machine vision. However, our study has a few limitations that need to be addressed in future research. For multispectral imaging, our study only examined a small part of the light spectrum. Additional equipment that can cover parts of the infrared and ultraviolet regions may aid in further advanced applications that need a wide range of wavelengths outside the visible light spectrum. Furthermore, we currently have not implemented cross-platform apps that are able to run the GUI needed for our applications, which may help when using our dome. This dome has immense potential as an educational and industrial tool to produce small object digital acquisition.

Acknowledgment. This work is supported by Universitas Gadjah Mada by Young Lecturer Grants 2022. This work is also supported by Electronics and Instrumentation Laboratory.

REFERENCES

- [1] R. J. Woodham, Photometric method for determining surface orientation from multiple images, *Optical Engineering*, vol.19, no.1, 191139, 1980.
- [2] S. Brenner, S. Zambanini and R. Sablatnig, An investigation of optimal light source setups for photometric stereo reconstruction of historical coins, *GCH2018 – Eurographics Workshop on Graphics and Cultural Heritage*, Vienna, 2018.
- [3] H. Yang and S. Hui, Analysis of light source selection and lighting technology in machine vision, *2021 International Conference on Computer Engineering and Application (ICCEA)*, pp.511-514, 2021.
- [4] A. Z. K. Frisky, A. Harjoko, L. Awaludin, S. Zambanini and R. Sablatnig, Investigation of single image depth prediction under different lighting conditions: A case study of ancient roman coins, *J. Comput. Cult. Herit.*, vol.14, 2021.
- [5] N. Hagen and M. Kudenov, Review of snapshot spectral imaging technologies, *Optical Engineering*, vol.52, 090901, 2013.
- [6] S. Dome and A. P. Sathe, Optical character recognition using tesseract and classification, *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, pp.153-158, 2021.
- [7] T. T. H. Nguyen, A. Jatowt, M. Coustaty and A. Doucet, Survey of post-OCR processing approaches, *ACM Comput. Surv.*, vol.54, 2021.
- [8] A. Harraj and N. Raissouni, OCR accuracy improvement on document images through a novel pre-processing approach, *Signal & Image Processing: An International Journal*, vol.6, 2015.
- [9] W. Gong, K. Zhang, C. Yang, M. Yi and J. Wu, Adaptive visual inspection method for transparent label defect detection of curved glass bottle, *2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL)*, pp.90-95, 2020.
- [10] H. Hayakawa, Photometric stereo under a light source with arbitrary motion, *Journal of the Optical Society of America A*, vol.11, no.11, pp.3079-3089, 1994.
- [11] L. W. MacDonald, E. Nocerino, S. Robson et al., 3D reconstruction in an illumination dome, *Proc. of the Conference on Electronic Visualisation and the Arts (EVA'18)*, pp.18-25, 2018.
- [12] L. W. MacDonald, Surface reconstruction from photometric normals with reference height measurements, *Proceedings of SPIE 9527, Optics for Arts, Architecture, and Archaeology V*, 952706, DOI: 10.1117/12.2184980, 2015.
- [13] L. MacDonald and S. Robson, Polynomial texture mapping and 3D representations, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives*, vol.38, pp.422-427, 2010.
- [14] L. MacDonald, Colour and directionality in surface reflectance, *AISB 2014 – The 50th Annual Convention of the AISB*, 2014.
- [15] Z. Zhou and P. Tan, Ring-light photometric stereo, in *Computer Vision – ECCV 2010. Lecture Notes in Computer Science*, K. Daniilidis, P. Maragos and N. Paragios (eds.), Berlin, Heidelberg, Springer, 2010.
- [16] A. McCarthy, K. Barton and L. Lewis, Low-cost multispectral imager, *Journal of Chemical Education*, vol.97, no.10, pp.3892-3898, 2020.
- [17] J. Park, E. Lee, Y. Kim, I. Kang, H. I. Koo and N. I. Cho, Multi-lingual optical character recognition system using the reinforcement learning of character segmenter, *IEEE Access*, vol.8, pp.174437-174448, 2020.
- [18] A. Mollah, N. Majumder, S. Basu and M. Nasipuri, Design of an optical character recognition system for camera-based handheld devices, *International Journal of Computer Science Issues*, vol.8, 2011.
- [19] X. Zhang, H. Ma, H. Zhao and R. Wang, Optical character recognition of electrical equipment nameplate with contrast enhancement, *2022 Power System and Green Energy Conference (PSGEC)*, pp.1062-1066, 2022.
- [20] R. Kaur and D. V. Sharma, Pre-processing for improving the recognition of mobile based Gurmukhi text recognition system, *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pp.1-5, 2021.
- [21] G. Vamvakas, B. Gatos, N. Stamatopoulos and S. Perantonis, A complete optical character recognition methodology for historical documents, *2008 8th IAPR International Workshop on Document Analysis Systems*, pp.525-532, 2008.

- [22] M. Brisinello, R. Grbić, D. Stefanovič and R. Pečkai-Kovač, Optical character recognition on images with colorful background, *2018 IEEE 8th International Conference on Consumer Electronics – Berlin (ICCE-Berlin)*, pp.1-6, 2018.
- [23] R. Yasir, M. Eramian, I. Stavness, S. Shirtliffe and H. Duddu, Data-driven multispectral image registration, *2018 15th Conference on Computer and Robot Vision (CRV)*, pp.230-237, 2018.
- [24] C.-H. Hung, T.-P. Wu, Y. Matsushita, L. Xu, J. Jia and C.-K. Tang, Photometric stereo in the wild, *2015 IEEE Winter Conference on Applications of Computer Vision*, pp.302-309, 2015.
- [25] C. Branas, F. Azcondo, R. Casanueva and F. Díaz, Optimal PWM implementation for dimming control of LED lamps, *PCIM Europe Conference Proceedings*, pp.1447-1454, 2014.
- [26] Y. Quéau, F. Lauze and J.-D. Durou, Solving uncalibrated photometric stereo using total variation, *J. Math. Imaging Vis.*, vol.52, pp.87-107, 2015.
- [27] V. Klema and A. Laub, The singular value decomposition: Its computation and some applications, *IEEE Transactions on Automatic Control*, vol.25, no.2, pp.164-176, 1980.
- [28] I. Jolliffe, *Principal Component Analysis*, John Wiley & Sons, Ltd., 2014.
- [29] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Transactions on Systems, Man, and Cybernetics*, vol.9, no.1, pp.62-66, 1979.