

Colorimetric Characterization of Color Image Sensors Based on Convolutional Neural Network Modeling

Po-Tong Wang,^{1*} Jui Jen Chou,^{1,2**} and Chiu Wang Tseng¹

¹Department of Bio-Industrial Mechatronics Engineering, National Taiwan University,
No. 1, Section 4, Roosevelt Road, Taipei 10617, Taiwan (R.O.C.)

²Department of Bio-Mechatronics Engineering, National Ilan University,
No. 1, Section 1, Shennong Road, Yilan City, Yilan County 26047, Taiwan (R.O.C.)

(Received March 13, 2018; accepted March 20, 2019)

Keywords: convolutional neural network, color image sensor, colorimetric characterization

The colorimetric characterization of a color image sensor was developed and modeled using a convolutional neural network (CNN), which is an innovative approach. Color image sensors can be incorporated into compact devices to detect the color of objects under a wide range of light sources and brightness. They should be colorimetrically characterized to be suitable for smart industrial colorimeters or light detection. Furthermore, color image sensors can be incorporated into machine vision systems for various industrial applications. However, the red, green, and blue (RGB) signals generated by a color image sensor are device-dependent, which means that different image sensors make different RGB spectrum responses under the same conditions. Moreover, the signals are not colorimetric; that is, output RGB signals are not directly coherent in terms of device-independent tristimulus values, such as CIE XYZ or CIELAB. In this study, the colorimetric mapping of RGB signals and CIELAB tristimulus values by CNN modeling was proposed. After digitalizing an RGB image sensor, characterizing RGB colors in the CIE color space, and CNN modeling for precise accuracy, the colorimetric characterization of color image sensors based on CNN modeling was proved to be superior to that based on $3 \times N$ polynomial regression. ΔE^*_{ab} was less than 0.5.

1. Introduction

Digital color cameras are widely used for capturing color images. The red, green, and blue (RGB) spectrum responses of CMOS image sensors⁽¹⁾ are neither colorimetrically nor linearly transformed from device-independent tristimulus values on the basis of CIE color-matching functions.⁽²⁾ State-of-the-art color-matching transform functions define the precision of color reproduction and mapping RGB signals to CIE XYZ or CIELAB. The transform derivation process is known as image sensor colorimetric characterization.

The International Organization for Standardization (ISO) has developed a standard (ISO 17321-4:2016)⁽³⁾ for digital camera color characterization, which is mainly used by camera manufacturers and testing laboratories. ISO requires the use of sophisticated, expensive

*Corresponding author: e-mail: neojwang@gmail.com

**Corresponding author: e-mail: jjchou@ntu.edu.tw

<https://doi.org/10.18494/SAM.2019.2271>

equipment and untended camera data. Therefore, the ISO 17321 standard is difficult to achieve in the color characterization and calibration of generic image-sensor-based devices. In this paper, a practical color-target-based method is proposed. The simplified method applies affordable IT 8.7 (ISO 1264) color calibration targets, which are captured by image sensors and measured using a spectrophotometer, to obtain RGB values and their corresponding XYZ values. Among color-matching transform methods such as polynomial regression,^(4,5) 3D lookup tables with interpolation and extrapolation,⁽⁶⁾ and various neural networks^(7,8) such as convolutional neural networks (CNNs)⁽⁹⁾ have inspired much academic and industrial research activities.^(10,11) However, no published research has used CNNs for the colorimetric characterization of image sensors.⁽¹²⁾ The deep learning of CNNs promises an intelligent solution for image sensor colorimetric characterization. A CNN is a deep machine learning algorithm to automatically learn, for example, color characterization features instead of features extracted by experts through polynomial regression. The CNN is a bioinspired trainable architecture that can learn invariant features from an input data set such as IT8.7/4 color samples. It also provides more accessible and smarter deep learning behaviors than trial-error manual testing by end users. Color science specialists can also deploy machine learning to adjust input vectors and optimize color-matching transform functions efficiently.

In this study, experiments were performed to investigate the following:

1. number of IT8.7/4 for CNN training data to build a model with high accuracy,
2. characterization performance when using CNNs,
3. characterization accuracy when using ΔE^*_{ab} as a loss function for CNN, and
4. comparison between CNN modeling and 3×11 polynomial regression.

2. Equipment and Materials

This study employed the following equipment and materials:

1. Printer and color reference targets: An EPSON Stylus Pro 9900 [Fig. 1(a)] was used to print out an adequate-quality test chart of IT8.7/4, 2005 Graphic technology-input data for the

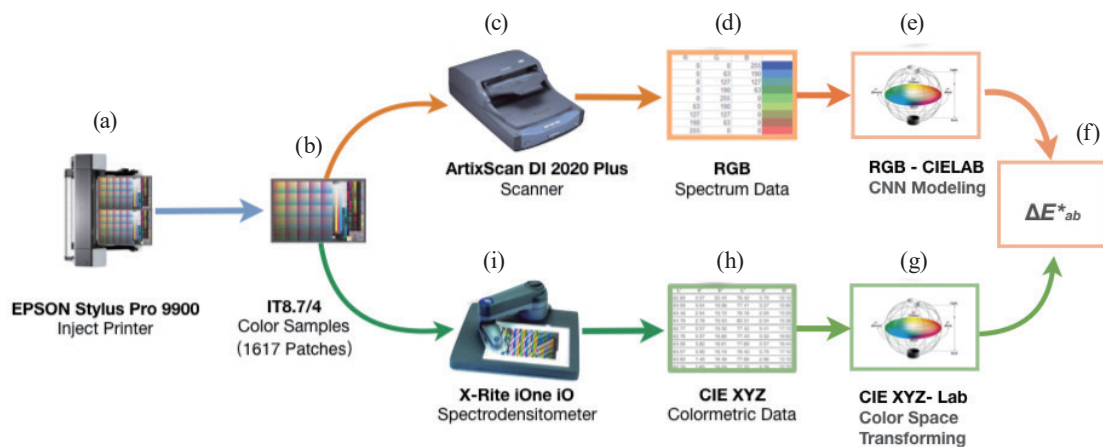


Fig. 1. (Color online) Equipment and workflow.

- characterization of 4-color process printing, 1617 color reference samples [Fig. 1(b)] on semiglossy photo paper (EasiColor).
2. Device for photographing RGB image data [Fig. 1(d)]: An ArtixScan DI 2020 Plus [Fig. 1(c)], which is a 3×8 -bit color scanner with both hardware and an optical resolution of 200×200 pixels. The digitally scanned data was directly transferred to a computer. The scanner was straightforward to set up and had adequate internal illumination for capturing the IT8.7/4 targets. An affordable image capturing device such as the ArtixScan DI 2020 Plus is suitable for academic testing and evaluation.
 3. Device for capturing CIE XYZ colorimetric data [Fig. 1(g)]: The X-Rite iOne iO spectrodensitometer [Fig. 1(f)] is an automatic chart-reading system designed to work with all iOnePro devices, which provide the fast measurement of reflective IT8.7/4 targets. Furthermore, CIELAB [Fig. 1(h)] quantities were defined and calculated using the CIE 1976 $L^*a^*b^*$ color difference formula [Fig. 1(i)].⁽¹³⁾
 4. The color characterization of the ArtixScan DI 2020 Plus scanner was performed by RGB-CIELAB color space transformation based on CNN modeling [Fig. 1(e)].
 5. CIELAB for consistency standard digital color image data (Fig. 2) for evaluating the characterization process: ISO 12640-3 Graphic technology—Prepress digital data exchange, Part 3: The CIELAB standard color image data (CIELAB/SCID) is applied to evaluate the performance of color pixel mapping between standard raw and CNN modeling [Fig. 1(e)].

3. CNN Characterization

Deep learning algorithms have equaled or surpassed human perception in some respects,^(14,15) enabling their deployment in real-world applications. In this study, a CNN algorithm was employed to achieve near-human perception in colorimetric characterization.



Fig. 2. (Color online) CIELAB standard color image data (CIELAB/SCID).

1. CNN training data:

(1) Input data of RGB spectrum data (RGB image) capturing:

- 1300 randomly selected color samples from IT8.7/4.
- The region of interest (ROI) of each color sample was composed of RGB colors of 8×8 pixels.

(2) Target data of CIELAB colorimetric data:

- Corresponded to the selected RGB data, by mapping the color space CIELAB of 1300 color samples from IT8.7/4.
- The ROI of each color sample was the color space CIELAB of 8×8 pixels.

2. CNN-validated data:

(1) Input data of RGB images:

- Remaining 317 color samples from IT8.7/4.
- The ROI of each color sample was composed of RGB colors of 8×8 pixels.

(2) Target CIELAB colorimetric data:

- Corresponded to the selected RGB data, by mapping the color space CIELAB of the remaining 317 color samples from IT8.7/4.
- The ROI of each color sample was the color space CIELAB of 8×8 pixels.

3. CNNs:

The CNNs used were trainable architectures composed of five stages (Fig. 3):

- $(8 \times 8 \times 3) - (8 \times 3 \times 8 \text{ C3}) - (8 \times 3 \times 16 \text{ C3}) - (8 \times 3 \times 32 \text{ C3}) - (8 \times 3 \times 64 \text{ C3}) - (8 \times 3 \times 128 \text{ C3}) - (8 \times 3 \times 256 \text{ FC}) - (\text{BP NN})$.
- $(8 \times 8 \times 3)$ is an 8×8 pixel RGB color sample of IT8.7/4.
- Convolution 3×3 (C3) is a 3×3 rectified linear unit (ReLU) convolution layer.
- No pooling layer was used.
- Flatten convolution (FC) is a fully connected layer that generated 6144 color characterizing features as input vectors of the backpropagation neural network (BPNN).

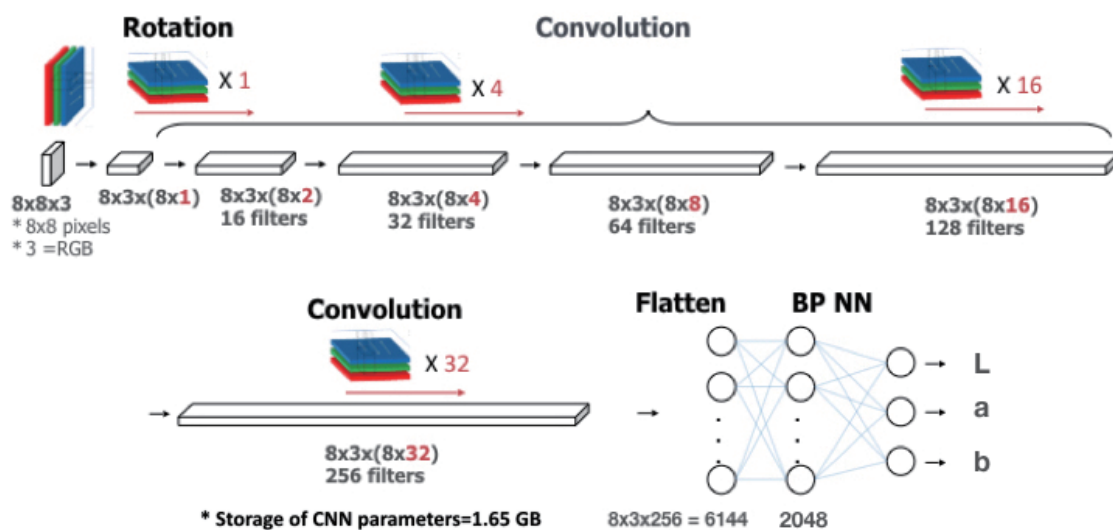


Fig. 3. (Color online) Color characterization using a CNN.

- Batch normalization with a minibatch of 1300 color samples was used to accelerate training.
 - A test error rate is associated with the expected validation error rate after 100000 training epochs.
4. 3×3 Mask:
 - (1) Initialized using a stochastic normal distribution.
 - (2) C3 is a 3×3 ReLU convolution layer, as shown in the 3×3 mask in Fig. 4.
 5. BPNN:
 - (1) Deep five-stage architectures, as shown in Figs. 3 and 4, were trained using supervised gradient backpropagation (BP) to ensure the accuracy of color characterization for image sensors. After training, the five-stage CNNs generated 6144 ($= 8 \times 3 \times 8 \times 32$) color features as input vectors of the BPNN. Then, a fully connected 2048 hidden layer was used to output color space CIE XYZ data.
 - (2) Loss function: In this study, we applied the CIE76 color difference formula ΔE_{ab}^* , which relates a regular set of coordinates to a known set of CIELAB coordinates as well as the loss function of the BPNN. The meaning of ΔE_{ab}^* that is approximately equal to 2.3 corresponds to a “just notable difference” compared with the human eye. A smaller ΔE_{ab}^* means that both color characterization and the neural network are optimized. The CIE76 color difference (ΔE_{ab}^*) formula is shown as

$$\Delta E_{ab}^* = \sqrt{(L_2^* - L_1^*)^2 + (a_2^* - a_1^*)^2 + (b_2^* - b_1^*)^2}. \quad (1)$$

In this study, a spectrodensitometer X-Rite iOne IO measured (L_1^*, a_1^*, b_1^*) and a BPNN inferred (L_2^*, a_2^*, b_2^*) ; these two colors were observed in the CIELAB color space, which expresses colors as three numerical values, namely, L^* for lightness and a^* and b^* for green–

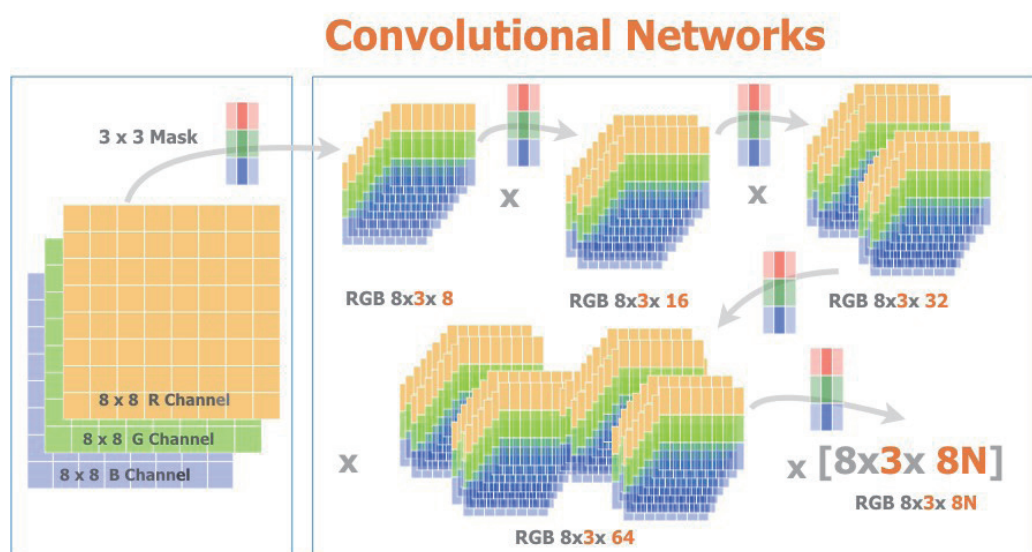


Fig. 4. (Color online) Convolutional networks.

red and blue–yellow color components. CIELAB was designed to be perceptually uniform concerning human color vision, meaning that the same amount of numerical change in these values corresponds to about the same amount of visually perceived difference. In the related works of this research, the loss function could be applied to other formulas for calculating ΔE —such as CMC and CIE94^(16,17)—for various industrial applications.

- (3) The loss function was also applied in adaptive moment estimation, the most widely used optimizer algorithm in CNN-related studies.

4. Results and Discussion

In this study, an average of 0.48 ΔE^*_{ab} was calculated using “Eq. (1)” [the equation of ΔE^*_{ab} shown as Eq. (1)]. The data substituted into the equation are the color target $L_1^*a_1^*b_1^*$ and the output value $L_2^*a_2^*b_2^*$ from the CNN model, respectively. The average of 0.48 ΔE^*_{ab} for the colorimetric characterization of an image sensor based on CNN modeling was superior to the average of 1 ΔE [units in CMC (1:1)] obtained when 3×11 polynomial regression was used.⁽⁴⁾ A functional validation using 1300 training and 317 validation color samples of IT8.7/4 by 3×11 polynomial regression obtained an average of 3.03 ΔE^*_{ab} and a maximum of 14.62 ΔE^*_{ab} . On the other hand, an average of 0.48 ΔE^*_{ab} and a peak of 11.49 ΔE^*_{ab} were obtained for CNN modeling. $\Delta E^*_{ab} > 1$ was achieved in 5.3% (85 color samples) of the IT8.7/4 1617 samples (Fig. 5), and $\Delta E^*_{ab} < 1$ was obtained in 94.7% (1532 color samples) of the IT8.7/4 1617 samples (Fig. 6). By comparing the CNN results, $\Delta E^*_{ab} > 1$ was obtained in 89.1% (1441 samples) of the IT8.7/4 1617 samples (Fig. 7), and $\Delta E^*_{ab} < 1$ was obtained in 10.9% (176 color samples) of the IT8.7/4 1617 samples (Fig. 8) for 3×11 polynomial regression.

There were four colors for which ΔE^*_{ab} exceeded 6, but the color difference was barely recognized by the human eyes (Fig. 9). The four colors with the highest ΔE^*_{ab} were as follows: color 1: $L(10), a(10), b(-3)$; CNN modeling: $L(10), a(-1), b(1)$; color 2: $L(21), a(-22), b(-11)$; CNN

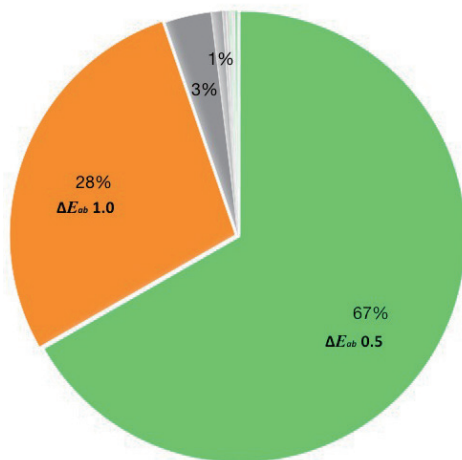


Fig. 5. (Color online) $\Delta E^*_{ab} > 1$ in 5.3% (85 samples) of the IT8.7/4 1617 samples (CNN modeling).

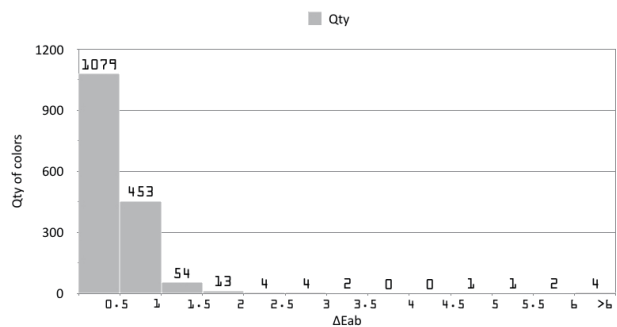


Fig. 6. $\Delta E^*_{ab} < 1$ in 94.7% (1532 samples) of the IT8.7/4 1617 samples (CNN modeling).

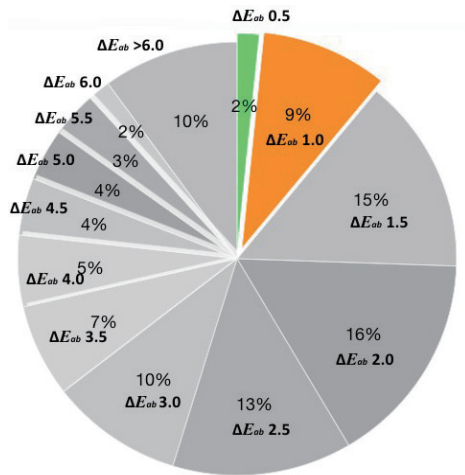


Fig. 7. (Color online) $\Delta E^*_{ab} > 1$ in 89.1% (1441 samples) of the IT8.7/4 1617 samples (3×11 polynomial regression).

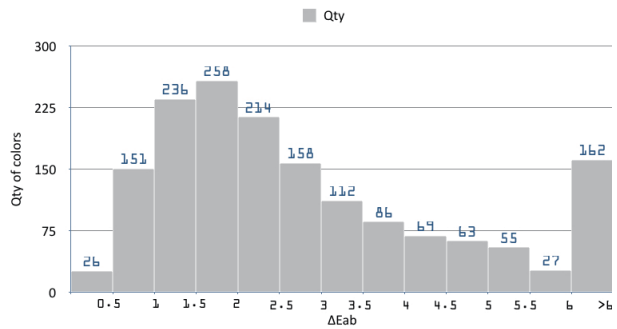


Fig. 8. $\Delta E^*_{ab} < 1$ in 10.9% (176 samples) of the IT8.7/4 1617 samples (3×11 polynomial regression).



Fig. 9. $\Delta E^*_{ab} > 6$ for four colors.

modeling: $L(24), a(-13), b(-9)$; color 3: $L(11), a(-9), b(2)$; CNN modeling: $L(10), a(-1), b(1)$; color 4: $L(10), a(8), b(0)$; CNN modeling: $L(10), a(-1), b(1)$.

We applied CIELAB-standard digital color image in Fig. 10 to visualize and evaluate color characterization. The CNN-characterized CIELAB image displayed in Fig. 11 and the 3×11 -polynomial-regression-characterized CIELAB image presented in Fig. 12 were compared with the original CIELAB image.

In summary, this paper has particular implications for the development of CNN algorithms that might offer an even smarter, more straightforward, and easier-to-use color reproduction system than $3 \times N$ polynomial regression. CNN, a new color space mapping technology, transforms a color-image-sensor-based device (i.e., camera) into a spectrometer-like device, which is reliable, capable, and responsive, and thus can serve as a foundation for future application.

However, some problems remain to be solved. The CNN model takes more than 100 h/1000 color patches on a PC with an Nvidia GPU GTX 1080 Ti. The validation of the CNN model



Fig. 10. (Color online) CIELAB-standard digital color image.



Fig. 11. (Color online) CNN-characterized CIELAB image.



Fig. 12. (Color online) 3×11 -polynomial-regression-characterized CIELAB image.

takes 68 ms for 8×8 pixels, making the method difficult to use in some real-time (<20 ms) industrial applications. Neural networks typically take longer to run as we increase the number of features, the number of hidden layers, and the number of columns in our dataset or image resolution. In this study, we run the CNN model on GPU, which primarily takes advantage of the parallel programming capabilities for mathematical operations. Training a model is only required once. During the stage of validation or testing, compared with the 3×11 polynomial regression of the color characterization validation, the CNN model takes 68 ms for 8×8 pixels, while 3×11 polynomial regression only needs less than 1 ms. 3×11 polynomial regression

seems to perform slightly faster than the CNN model. However, the accuracy of CNN modeling is much higher than that of polynomial regression. The processing time may not satisfy some industrial applications. However, we believe that the processing power of computer hardware would progress rapidly as Moore's law indicated that "the power of computers would double every 12 months, while the cost of that technology would fall by 50% over the same time".

Moreover, IT8.7/4 only provides 1617 color reference samples for color patch training. It might lead to a CNN model that is slightly overfitting but does not markedly affect the CNN performance. In our future work, we propose to capture more training data of color patches and use data augmentation to reduce overfitting. On the other hand, we will use the dropout mechanism and batch normalization of CNN programs to possibly prevent overfitting. To speed up CNN computing, optimized CNN architectures, distributed CNN computing, and FPGA-based CNN accelerators for meeting real-world applications will be used.

5. Conclusions

In this study, modeling for the colorimetric characterization of a color image sensor was developed and validated using a CNN. CNN represents a new method of extracting color features automatically. The high accuracy of CNN characterization was demonstrated for 8×8 pixel color image sequential characterization without requiring any greyscale balancing or image preprocessing. A satisfactory result of a mean color difference ΔE^*_{ab} of as small as 0.48 was obtained.

References

- 1 B. Choi, M. Bae, S. Kim, J. Lee, C. Oh, S. Seo, S. Chang, J. Park, and J. Shin: *Sens. Mater.* **29** (2017) 235.
- 2 CIE, Colorimetry, CIE Pb. 15:2004 (CIE Central Bureau, Vienna, 2004).
- 3 ISO/TS 17321-4:2016 1732, Graphic technology and photography Color characterization of digital still cameras (DSCs) Part 4: Programmable light emission system (2016).
- 4 G. Hong, M. R. Luo, and P. A Rhodes: *Color Res. Appl.* **26** (2001) 76.
- 5 K. León, D. Mery, F. Pedreschi, and J. León: *Food Res. Int.* **39** (2006) 1084.
- 6 P. C. Hung: *J. Electron. Imaging* **2** (1993) 53.
- 7 H. R. Kang and P.G. Anderson: *J. Electron. Imaging* **1** (1992) 125.
- 8 T. L.V. Cheung and S. Westland: *CIC* **4** (2002) 117.
- 9 F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer: *cs.CV* 4 arXiv: 1602.07360. (2016).
- 10 A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam: *cs.CV* 1 arXiv: 1704.0486. (2017).
- 11 S. Ioffe and C. Szegedy: *PMLR* **37** (2015) 448.
- 12 Y. F. Chou, M. R. Luo, C. Li, V. Cheung, and S. L. Lee: *Color. Technol.* **129** (2012) 203.
- 13 A. R. Robertson: *Color Res. Appl.* **15** (1990) 167.
- 14 A. Krizhevsky, I. Sutskever, and G. Hinton: *NIPS* **1** (2012) 1097.
- 15 G. Li, P. Li, G. Sawaya, G. Gopalakrishnan, I. Ghosh, and S. P. Rajan: *Proc. 17th ACM SIGPLAN Symp. Principles and Practice of Parallel Programming (PPoPP '12)* (ACM, New York, 2012) 215.
- 16 F. J. J. Clarke, R. McDonald, and B. Rigg: *J. Soc. Dyers Colour.* **100** (1984) 128.
- 17 CIE Pub. 116-1995, Industrial Colour-difference Evaluation (CIE Central Bureau, Vienna, 1995).

About the Authors



Po-Tong Wang received his M.S. degree in Computer Science from National Taipei University of Education, Taiwan, in 2015. He is a Ph.D. student supervised by Dr. Jui Jen Chou in the Bio-Industrial Mechatronics Engineering Department at NTU, Taiwan since 2016. His primary research interests include image sensors, color space transformation algorithms, and PLCopen control for advanced industrial automation and industrial applications of image vision and automation in Taiwan. (neojwang@gmail.com)



Jui Jen Chou received his Ph.D. (1990) and M.S. (1986) degrees in Mechanical Engineering from UCLA, and his M.S. (1984) and B.S. (1980) degrees in Agricultural Engineering from National Taiwan University (NTU), Taiwan. His major research interests include mechatronics, biorobotics, bioindustry, automation, and engineering education. He has been the Vice President for Academic Affairs and the Chair Professor of National Ilan University, Taiwan since 2016. He has been a distinguished professor of the Bioindustrial Mechatronics Engineering Department at NTU, Taiwan since 2008. He was the Deputy Vice President for Academic Affairs, NTU (2013-2016), the Director of Center for Teaching and Learning Development, NTU (2013-2016); the Chairman of Bioindustrial Mechatronics Engineering Department, NTU (2009-2012), the Director of the Education and Research Center for Agriculture Automation, NTU (1998-2004); the President of Taiwan Institute of Biomechatronics (2012-2015). He received numerous awards, including outstanding teaching awards from NTU, an outstanding academic achievement award from Chinese Institute of Agricultural Machinery, and an outstanding academic achievement award from Taiwan Institute of Biomechatronics, and so forth. (jjchou@ntu.edu.tw)



Chiu Wang Tseng received his M.S. degree in the Department of the Bio-Industrial Mechatronics Engineering from National Taiwan University (NTU), Taiwan, in 2018. His primary research interests include deep learning text mining and mechatronics. (he01181603@gmail.com)