

A Study on Chromaticity Modulation Methods Based on Pixel-Level Color Difference in Display-to-Camera Communication

Kosuke Iwasaki

Dept. Electrical and Electronic Engineering

Meijo University

Nagoya, Japan

263427008@ccmail.meijo-u.ac.jp

Kentaro Kobayashi

Dept. Electrical and Electronic Engineering

Meijo University

Nagoya, Japan

kkobaysh@meijo-u.ac.jp

Abstract—This paper investigates chromaticity modulation methods for display-to-camera visible light communication using a digital signage display and an image sensor. Conventional uniform-intensity methods cause locally perceptible color changes because signal perceptibility depends on pixel color. To address this problem, this paper proposes a pixel-level chromaticity modulation method that determines the signal intensity for each pixel based on color difference. A normalized cross-correlation (NCC)-based demodulation method using the signal-intensity distribution within each cell is also investigated. Experimental results show that the proposed modulation method improves the uniformity of visual quality and reduces the bit error rate under many conditions.

Keywords—Visible light communication, Digital signage, Image sensor, YCbCr color space, Color difference

I. INTRODUCTION

Digital signage is a general term for media that disseminate information using electronic display devices. It has rapidly become widespread and is now installed in various locations, such as train stations, shopping malls, and hospitals. This widespread use is mainly due to the ease of updating displayed content and the ability to provide information according to the installation location and time of day. In this context, visible light communication (VLC) systems have been investigated to add information-transmission functionality to image content and to allow mobile terminals to receive data. In such a system, the transmitter embeds data signals by changing the luminance or chromaticity components of image content displayed on digital signage, and the receiver recovers the information by capturing and analyzing the displayed image with an image sensor in a mobile terminal. This enables users to receive real-time value-added information, such as coupons, event details, and disaster guidance, together with visual advertisement content.

Several display-to-camera VLC methods have been reported [1, 2], including marker-based methods with imperceptible switching speed (e.g., [3]) and high-speed backlight modulation (e.g., [4]); however, these approaches suffer from an extremely low data rate (about 10 bps) or require hardware modification, respectively. Software-based spatial parallel transmission methods have also been investigated, including frequency-domain modulation [5, 6] and modulation using chromaticity components to which the human eye is less sensitive [7, 8]. Chromaticity-based

modulation has the advantage of causing little degradation in visual quality. However, because conventional methods apply a uniform signal intensity to all pixels, the perceptibility of the embedded signal varies depending on the pixel color. As a result, there can occur local regions in which the embedded signal becomes perceptible.

This paper proposes a pixel-level chromaticity modulation method that determines the signal intensity for each pixel individually and evaluates it against the conventional uniform-intensity method in terms of both visual and communication quality.

II. SYSTEM MODEL

We consider a visible light communication system in which the transmitter is a display and the receiver is a camera. Fig. 1 shows the system model, and Fig. 2 shows the frame-difference acquisition method.

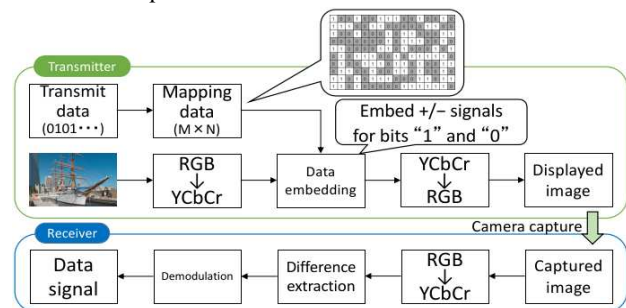


Fig. 1. System model.

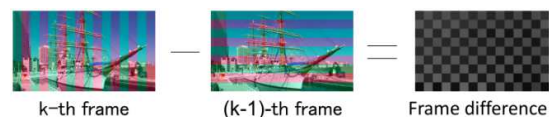


Fig. 2. Frame-difference acquisition method.

A. Modulation

In the modulation process, the input image is first converted from the RGB color space to the YCbCr color space. This conversion enables data embedding in either the Cr component, which corresponds to the red–green chromaticity axis, or the Cb component, which corresponds to the yellow–blue chromaticity axis. This approach exploits the fact that the human visual system is less sensitive to chromaticity changes than to luminance changes. The transmit data sequence is a binary symbol sequence consisting of 0 and 1. The display area is divided into $M \times N$

A part of this research was supported by JSPS KAKENHI Grant Number 23H01413, 23K26108.

cells, and the mapped data corresponding to each cell are added to either the Cr or Cb component of the image. Differential encoding is applied so that data are represented by signal changes between two consecutive frames. This encoding is used because the receiver takes inter-frame differences to remove the image content component in the demodulation process.

Baseband binary phase-shift keying (BPSK) is used as the embedding scheme. The signal value embedded in each frame is set to $\pm\alpha/2$, where α denotes the signal intensity. When the transmit bit is 1, the signal polarity in the k -th frame is inverted relative to that in the $(k-1)$ -th frame, namely from $+\alpha/2$ to $-\alpha/2$ or vice versa. When the transmit bit is 0, the polarity is maintained. Through this differential encoding, data are embedded in the image as temporal changes. Finally, the signal-embedded image is converted back from the YCbCr color space to the RGB color space to generate the displayed image.

In this study, the signal intensity is determined using the CIEDE2000 color-difference formula [9], which accounts for human color-vision and is used as a quantitative metric for visual quality. CIEDE2000 is defined by the International Commission on Illumination (CIE), and a color difference of 1.0 or less is generally regarded as difficult to perceive. The signal intensity α is determined as follows.

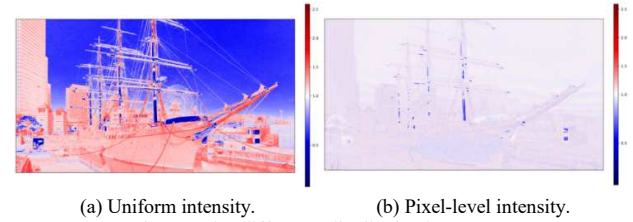
In the conventional uniform-intensity chromaticity modulation method used in the previous work [8], a single signal intensity is applied to the entire image. Specifically, the root mean square (RMS) of the color differences over all pixels is calculated using two frames in which $+\alpha/2$ and $-\alpha/2$ are embedded, respectively. The value of α is then determined so that the RMS equals the target color difference, and the single value is used as the signal intensity for each image. This method, however, can produce regions in which the color difference is excessively large and regions in which it is insufficient, leading to degradation in both visual quality and communication quality.

To address this limitation, this study proposes a chromaticity modulation method with pixel-level intensity optimization. In the proposed method, the color difference is calculated for each pixel in the same manner as in the conventional method, and the signal intensity is set so that each pixel directly satisfies the target color difference. For a 4K image with a resolution of 3840×2160 pixels, the signal intensity must be determined for approximately 8.3 million pixels. To find the signal intensity that satisfies the target color difference for each pixel, a binary search algorithm is applied. The search range for the added chromaticity value is set to $[0.0001, 0.01]$, and the range is narrowed according to whether the calculated color difference is larger or smaller than the target value. This process enables efficient calculation of the signal intensity for each pixel.

In both the conventional and proposed methods, the color difference is evaluated over all pixels when determining the signal intensity. The conventional method uses this evaluation to compute the RMS color difference over the entire image, whereas the proposed method uses the evaluation to update the signal-intensity map for individual pixels. In the proposed method, the pixel-wise signal intensities are updated for all pixels in each binary-search

iteration. Therefore, for an image with P pixels and I search iterations, the computational cost of the search is proportional to PI . In this study, the maximum number of search iterations is set to 20. Although the proposed method requires pixel-wise signal-intensity maps and increases the memory requirement, this optimization is performed only once at the transmitter side when generating the signal-embedded content. Therefore, it does not increase the computational load of real-time demodulation at the receiver. The proposed method can achieve uniform visual quality over the entire image.

Fig. 3 shows a heatmap example of the per-pixel color difference when Cb-component modulation is applied with a target color difference of 1.0. The scale ranges from 0.07 to 2.59. As shown in the heatmap, the conventional method produces regions in which the color difference is excessively large, indicated in red, and regions in which it is too small, indicated in blue. In contrast, the proposed method maintains the color difference near the target value over the entire image.



(a) Uniform intensity. (b) Pixel-level intensity.
Fig. 3. Color-difference distribution heatmaps.

B. Demodulation

In the demodulation process, the differentially encoded signal is extracted from the image captured by the camera, and the transmitted data are recovered. First, the captured image is converted from the RGB color space to the YCbCr color space, and the chromaticity component in which the data are embedded is extracted. Next, the inter-frame difference between consecutive frames is calculated to remove the image content component. Based on the extracted differential signal, this study investigates two alternative data-decision methods to reconstruct the transmitted bits: (1) a cell-average-based decision and (2) a normalized cross-correlation (NCC)-based decision.

1) Cell-Average-Based Decision:

For the extracted signal, the absolute value of the mean pixel value in each cell is calculated, and a binary decision for bit 0 or 1 is made using a threshold. For each cell, the mean pixel value is calculated in each of six pilot frames transmitted in advance, and the threshold is set to the average of these six values. Three pilot frames correspond to bit 0, and the other three correspond to bit 1.

2) NCC-Based Decision:

In the pixel-level intensity method, the signal intensity has a spatial distribution within each cell. Therefore, a decision method that accounts for this signal-intensity distribution can be expected to improve demodulation performance. This study investigates a demodulation method based on NCC, which utilizes spatial pattern information within each cell. During demodulation, the NCC score $S_{m,n}$ for each cell is calculated between the data frame after the pilot frames and the estimated pattern $\beta_{x,y}$, based on which a binary decision is made. The NCC score is defined as:

$$S_{m,n} = \frac{\sum_{x,y \in C_{m,n}} d_{x,y} \cdot \beta_{x,y}}{\sqrt{\sum_{x,y \in C_{m,n}} d_{x,y}^2 \cdot \sum_{x,y \in C_{m,n}} \beta_{x,y}^2}}$$

Here, $d_{x,y}$ denotes the received signal intensity of each pixel obtained from the inter-frame difference, $\beta_{x,y}$ denotes the estimated pattern generated from the pilot frames, and $\sum_{x,y \in C_{m,n}}$ denotes the summation over all pixels in cell (m,n) .

The estimated pattern $\beta_{x,y}$ and the decision threshold are determined using six pilot frames transmitted in advance. These consist of three pilot frames with signals corresponding to bit 1 and three pilot frames with signals corresponding to bit 0. The estimated pattern $\beta_{x,y}$ is generated as the pixel-wise average of the three pilot frames corresponding to bit 1.

Let $S_{m,n}^{(1)}$ denote the average NCC score between each bit-“1” pilot frame and the pixel-wise average of the remaining two bit-“1” pilot frames. Similarly, let $S_{m,n}^{(0)}$ denote the average NCC score between each bit-“0” pilot frame and the pixel-wise average of the remaining two bit-“0” pilot frames. The decision threshold $\theta_{m,n}$ is defined as

$$\theta_{m,n} = (S_{m,n}^{(1)} + S_{m,n}^{(0)})/2$$

During demodulation, the NCC score $S_{m,n}$ is calculated between the data frame and the estimated pattern $\beta_{x,y}$. The received bit is decoded as “1” if $S_{m,n} > \theta_{m,n}$; otherwise, it is decoded as “0.”

III. EXPERIMENTAL RESULTS

A. Evaluation Setup

The transmitter display is an LG Electronics OLED55C5PJA. The receiver camera is a DFK37BUX290 from The Imaging Source mounted on a tripod and equipped with an 8-mm lens. The experimental parameters for the communication-quality evaluation are summarized in Table I. Because the objective of the experiments is to evaluate both communication quality and visual quality, both the transmitter and receiver are kept stationary. Fig. 4 shows the experimental setup. Ten images from the Ultra-high definition/wide-color-gamut standard test images provided by the Institute of Image Information and Television Engineers [10], shown in Fig. 5, are used as the signage images into which data signals are embedded. To focus on changes in communication quality caused by differences in the modulation methods, the display image update and camera capture are synchronized. Thus, errors caused by synchronization mismatch between the display and camera are excluded from the evaluation.

The bit error rate (BER) calculated by comparing the received and transmitted data and counting erroneous bits, using the same data sequence for all transmissions, is used as the metric for communication reliability. To reduce the influence of environmental disturbances during image capture, such as sunlight and shadows cast by pedestrians, the BER is averaged over 350 frames, corresponding to 50,400 bits, for each chromaticity component and each target color difference. The target color difference is varied from 0.4 to 1.0 to evaluate its effect on communication quality, with 1.0 used as the reference value because color differences of 1.0 or less are generally difficult for the human eye to perceive.

TABLE I. EXPERIMENTAL PARAMETERS

Transmitter	
Resolution	3840×2160 [pixel]
Cell configuration (M, N)	(9, 16)
Color temperature of white point	6500 [K]

Receiver	
Resolution	1920×1080 [pixel]
White balance	6500 [K]
Exposure time	1/60 [s]
Distance	3.0 [m]



Fig. 4. Experimental setup.



Fig. 5. Ultra-high definition/wide-color-gamut standard test images.

B. Visual Quality Evaluation

Figs. 6 and 7 show histograms of the per-pixel color difference aggregated over all ten standard images when the target color difference is set to 1.0, for Cr- and Cb-component modulations, respectively. Table II summarizes the statistical results for each method. The pixel-level intensity method achieves a much narrower color-difference distribution centered around the target value than the conventional uniform-intensity method. Compared with the uniform-intensity method, the proposed method reduces the standard deviation by 70.7% for the Cr component and by 63.9% for the Cb component, as summarized in Table II.

The concentration of the histogram around the target value indicates that the visual quality is more uniform over the entire image. In the conventional method, the broad color-difference distribution means the image-wide average can satisfy the target value while some pixels still show large, perceptible color differences. In contrast, the proposed method keeps the color differences of all pixels close to the target value, reducing such perceptible changes.

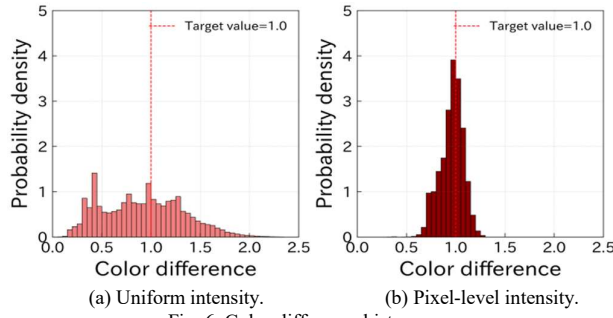


Fig. 6. Color-difference histograms.
(10 images combined, Cr modulation, target $\Delta E = 1.0$).

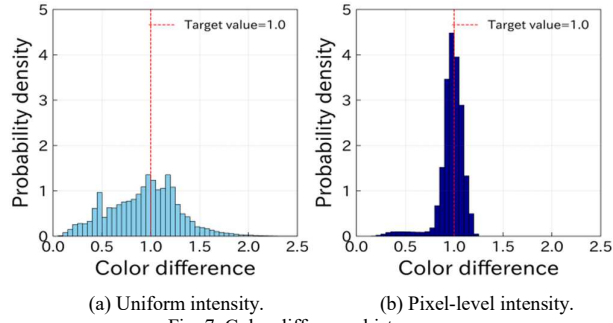


Fig. 7. Color-difference histograms.
(10 images combined, Cb modulation, target $\Delta E = 1.0$).

TABLE II. VISUAL QUALITY COMPARISON

Component	Method	RMS	Std.Dev.	Q1	Q3	IQR
Cr	Uniform	1.00	0.41	0.56	1.21	0.65
	Pixel-level	0.97	0.12	0.89	1.04	0.15
Cb	Uniform	1.00	0.36	0.68	1.17	0.49
	Pixel-level	0.98	0.13	0.93	1.05	0.12

C. Communication Quality Evaluation

Fig. 8 shows the BER results for the ten standard images under three conditions: the conventional method, which uses uniform-intensity modulation and cell-average-based demodulation; the proposed pixel-level-intensity modulation with cell-average-based demodulation; and the proposed modulation with NCC-based demodulation. In each condition, the signal intensity corresponding to each target color difference is used. When no bit errors are detected during the transmission of the 350 frames, the transmission is regarded as error-free and is plotted as a BER of 10^{-5} .

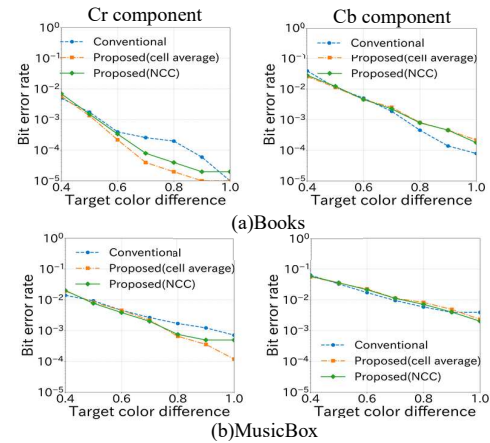
The pixel-level-intensity method reduces the BER under most conditions for the ten images and for both the Cr and Cb components, as shown in Fig. 8. This improvement is attributed to the increase in the average signal intensity over the image. In the uniform-intensity method, the color difference is calculated as the RMS over the entire image. Therefore, when the color-difference variation is large, the signal intensity must be reduced to satisfy the target color difference. In contrast, the proposed method determines the signal intensity independently for each pixel, thereby increasing the average signal intensity over the image and reducing the BER.

To further investigate the relationship between signal intensity and BER, all 140 test conditions were evaluated, consisting of ten images, two chromaticity components (Cr and Cb), and seven target color differences (0.4–1.0).

Compared with the conventional method, the proposed pixel-level method with cell-average-based decision improved the BER in 78 conditions, produced no change in 18 conditions, and degraded the BER in 44 conditions; the median BER improvement rate across the 140 conditions was 18.6%.

To clarify the cause of these condition-dependent improvements and degradations, the average signal intensities of the cells in which decoding errors occurred were compared. In the improved conditions, when a cell was decoded incorrectly by the conventional method but correctly by the proposed method, the average signal intensity of the proposed method at that cell was 1.68 times higher than that of the conventional method. Conversely, in the degraded conditions, when a cell was decoded correctly by the conventional method but incorrectly by the proposed method, the average signal intensity of the proposed method at that cell was only 0.88 times that of the conventional method. These results indicate that the improvement or degradation in BER is closely associated with an increase or decrease in the per-cell average signal intensity, confirming that the communication-quality gain of the proposed method is driven primarily by the increase in average embedded signal intensity rather than by a qualitative change in the demodulation process itself.

The NCC-based decision was introduced to exploit the spatial distribution of signal intensity within each cell, and was therefore expected to outperform the simpler cell-average-based decision, which discards this spatial information. In practice, however, the NCC-based decision provided little additional improvement in BER, and the final BERs of the two demodulation methods were nearly identical. This result can be attributed to the sensitivity of NCC-based pattern matching to small positional displacements and noise, to which the cell-average-based decision, aggregating the signal over many pixels, is inherently more tolerant. Fig. 9 compares the temporal variation in the cell-average difference with that at a single center pixel in cell (6,4), when bits 1 and 0 are transmitted alternately, using the ChromaKey image, Cr-component modulation, and a target color difference of 0.7. As shown in the figure, the cell-average difference follows the embedded signal, alternating between large and small values, whereas the center-pixel value partially fails to follow this transition.



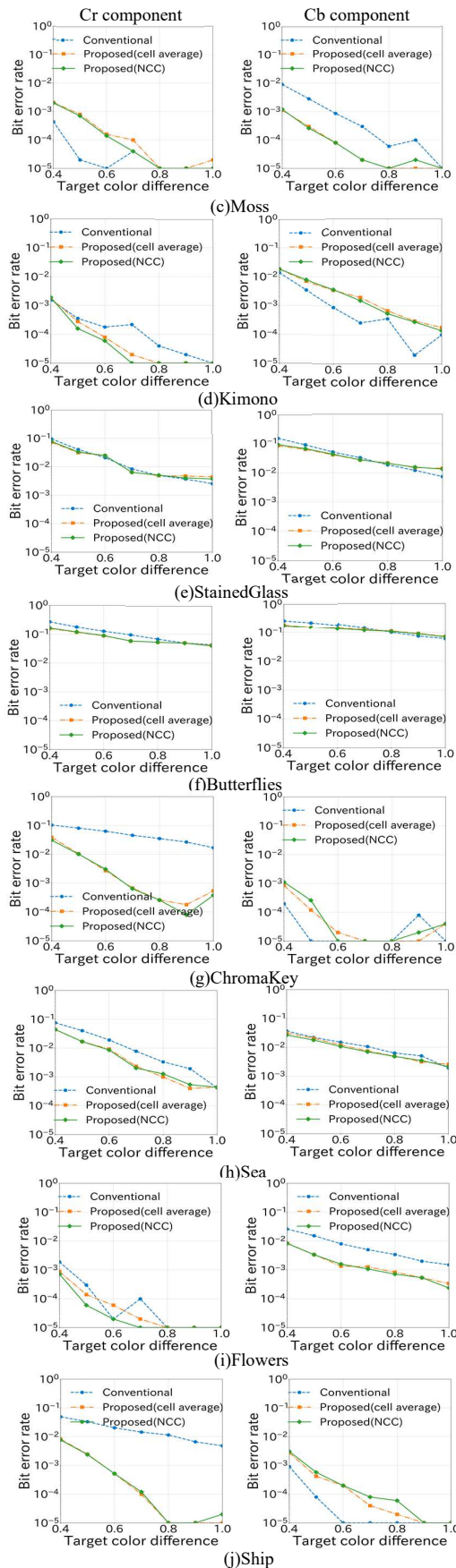


Fig. 8. Comparison of BERs for each image.

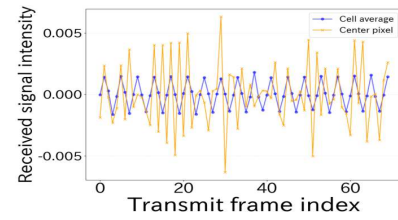


Fig. 9. Received signal intensity transitions.

IV. CONCLUSION

This paper proposed a pixel-level chromaticity modulation method for display-to-camera visible light communication that determines the signal intensity for each pixel to satisfy the target color difference, together with an NCC-based demodulation method that uses the signal intensity distribution within each cell.

The visual quality evaluation showed that the proposed method concentrated the color differences near the target value and reduced the standard deviation by 70.7% for the Cr component and 63.9% for the Cb component compared with the uniform-intensity method. The communication quality evaluation showed that the proposed method reduced the BER under many conditions by increasing the average signal intensity. The NCC-based decision showed performance comparable to the cell-average-based decision because of its sensitivity to slight positional displacement and noise. Evaluation under practical conditions, including camera motion, synchronization mismatch, perspective distortion, and lighting variation, remains for future work.

REFERENCES

- [1] M. Kinoshita, Y. Kozawa, K. Kobayashi, S. Arai, and T. Yamazato, "Survey of Visible Light Communication Research in Japan," *IEICE Trans. Fundamentals*, art. no. 2025EAP1194, Jan. 2026.
- [2] P. Singh, Y.J. Kim, B.W. Kim, and S.Y. Jung, "Display field communication: Enabling seamless data exchange in screen-camera environments," *Photonics*, vol. 11, no. 11, p. 1000, Oct. 2024.
- [3] K. Kuraki, S. Nakagata, R. Tanaka, and T. Anam, "Video-mediated communication technology for display-to-mobile-terminal communication," *FUJITSU*, vol. 64, no. 5, pp. 504–509, Sep. 2013.
- [4] H. Aoyama and M. Oshima, "Visible light communication using a conventional image sensor," *Proc. IEEE Consumer Commun. Netw. Conf. (CCNC)*, pp. 103–108, Jan. 2015.
- [5] Y. Lin, T. Wada, K. Mukumoto, and H. Okada, "Performance evaluation of information superimposing schemes based on wavelet transform for parallel transmission visible light communication systems," *Proc. IEEE Global Conf. Consumer Electron. (GCCE)*, Oct. 2017.
- [6] R. Mushu, T. Wada, K. Mukumoto, and H. Okada, "A Proposal of Information Superimposing Scheme Based on Discrete Cosine Transform in Parallel Transmission Visible Light Communications," *Proc. IEEE Global Conf. Consumer Electron. (GCCE)*, Oct. 2018.
- [7] H. Okada, S. Sato, T. Wada, K. Kobayashi, and M. Katayama, "Preventing degradation of the quality of visual information in digital signage and image-sensor-based visible light communication systems," *IEEE Photonics J.*, vol. 10, no. 3, Art. no. 7903509, Apr. 2018.
- [8] K. Shimei, K. Kobayashi, and W. Chujo, "Effect of perceptually uniform color space and diversity of chromaticity components on digital signage and image sensor-based visible light communication," *IEICE Trans. Fundamentals*, vol. E107-A, no. 4, pp. 638–653, Aug. 2023.
- [9] M. R. Luo, G. Cui, and B. Rigg, "The development of the CIE 2000 colour-difference formula: CIEDE2000," *Color Res. Appl.*, vol. 26, no. 5, pp. 340–350, Oct. 2001.
- [10] Institute of Image Information and Television Engineers, Ultra-high definition/wide-color-gamut standard test images, <https://www.ite.or.jp/content/chart/uhdvtv/>