Overlaying a motion video with data information for digital signage and imagesensor-based visible light communication systems*

Hiraku Okada¹a), Shota Yoshida², Tadahiro Wada³, Kentaro Kobayashi¹, and Masaaki Katayama¹

¹ Institute of Materials and Systems for Sustainability, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464–8603, Japan
² Dept. of Information and Communication Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464–8603, Japan
³ Graduate School of Integrated Science and Technology, Shizuoka University, Johoku 3–5–1, Naka-ku, Hamamatsu 432–8561, Japan

a) hiraku@m.ieice.org

Abstract: A digital signage image-sensor-based visible light communication (IS-VLC) system can transmit data information by embedding it on visual information. The use of a motion video as visual information makes it difficult to completely erase the visual information in this system. The extraction of data information is an important issue because the incomplete extraction will degrade the bit error rate performance. In this letter, we expand the IS-VLC system to overlay a motion video with data information by introducing a cell-position correction method and a noise elimination method. Experimental results clarify the effectiveness of the extended system.

Keywords: digital signage, visible light communication, image sensor, motion video

Classification: Wireless Communication Technologies

References

[1] S. Sato, H. Okada, K. Kobayashi, T. Yamazato, and M. Katayama, “Visible light communication systems using blue color difference modulation for digital signage,” IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 1242–1247, 2016. DOI:10.1109/PIMRC.2016.7794761

[2] S. Yoshida, H. Okada, T. Wada, K. Kobayashi, and M. Katayama, “Performance evaluation of digital signage and image-sensor-based visible light communication systems using motion video as visual information,” Workshop on Smart City based on Ambient Intelligence (SCAI), 2018.

*An earlier versions of this letter were presented at IEEE International Symposium on Personal, Indoor and Mobile Radio Communications [1] and Workshop on Smart City based on Ambient Intelligence [2].
1 Introduction

Digital signage is a display that is used to display targeted information, entertainment, advertisements, and announcements. Digital signage is placed in stations, shopping malls, and other such places. The dissemination of accurate emergency information during disasters and a guide for foreign passengers [3, 4] are advanced usage of digital signage. These applications require not only a display of visual information but also transmissions of data information from the digital signage to user equipment such as smartphones.

We focus on an image-sensor-based visible light communication (IS-VLC) system [5, 6] for the transmission of data information. In this letter, the IS-VLC system employs a display of a digital signage as a transmitter and an image sensor
as a receiver. The transmitter overlays visual information with data information, and the receiver derives data information from the frames captured by the image sensor.

In the IS-VLC system, it is important to prevent quality degradation of visual information that is overlaid with data information. There are schemes to solve this issue [7, 8, 9], but these schemes have problems such as a low data rate or requirement of a special display. To resolve these problems, we proposed color differential modulation for the IS-VLC system and clarified the capability of the transmission of data information while avoiding the quality degradation of the visual information [10]. However, still images are used as visual information. When a motion video is used as visual information, it is difficult to completely erase visual information. The extraction of data information is an important issue because the incomplete extraction will degrade the bit error rate (BER) performance.

Herein, we expand the IS-VLC system to overlay a motion video with data information, which introduces a cell-position correction method and a noise elimination method. We evaluate the BER performance of the expanded system and clarify the effectiveness of these methods.

2 IS-VLC system

Fig. 1 shows the model of the expanded IS-VLC system. In this section, we explain the conventional IS-VLC system proposed in [10]. The cell-position correction method and the noise elimination method is described in Sect. 3.

At the transmitter, bit sequence of data information is mapped onto an $M \times N$ data matrix, which is differentially encoded to obtain a coded matrix. A mono-component data frame is generated from the coded matrix, where the size of the data frame is $X \times Y$ pixels. The data frame consists of $M \times N$ rectangular cells, and the size of the cell is $X/N \times Y/M$ pixels. Each cell corresponds to a bit of the coded matrix. The pixel value of the cell is 0 and $\alpha$ when the corresponding bit is 0 and 1,
respectively, where $\alpha$ indicates a signal intensity. The color-difference modulator overlays blue-color-difference (Cb) components of visual information with the data frame because the human eyes are insensitive to color-difference components, especially the Cb component. After adding a marker, which is used at the receiver, the overlaid frame is shown on a display.

At the receiver, an image sensor captures the frame on the display. Using the marker, the region of the data frame is extracted from the captured frame, and time-synchronization is achieved. Through a color filter, a Cb-component frame is derived, and the data frame is retrieved by subtracting the successive two frames. This subtraction can decode the data differentially encoded by the transmitter. The pixel values in each cell are averaged, and the data matrix is retrieved via a threshold decision. Finally, the bit sequence of the data information is derived by demapping the retrieved data matrix.

When a still image is used as visual information, the subtraction eliminates visual information from the received frame. However, when a motion video is used, visual information cannot be eliminated completely, and the residual values of visual information will degrade the BER performance.

### 3 Expansion for motion video

In this section, we introduce a cell-position correction method and a noise elimination method for a motion video as visual information.

#### 3.1 Cell-position correction

When the motion video gradually changes, the movement between successive frames will be small. We introduce a cell-position correction method based on template matching, which will be able to reduce the residual values after the subtraction by adjusting the position of the successive frames.

The cell-position correction is achieved at each cell. Let $C_{m,n}(i)$ show the cell of the $m$th row and the $n$th column in the $i$th frame. First, the cell $C_{m,n}(i - 1)$ is used as a template. For the $i$th frame, this method searches and finds the position at which the similarity between the cell $C_{m,n}(i)$ and the template is the largest. In this process, only luminance (Y) components of both cells are used because Y components do not include data information. Herein, we employ a normalized cross-correlation as similarity [11]. Following the color filter for Cb components, only the duplicated regions between the successive cells $C_{m,n}(i - 1)$ and $C_{m,n}(i)$ are subtracted after adjusting their positions so that the similarity takes the largest value.

#### 3.2 Noise elimination

The change between the successive frames of the motion video causes residual values of visual information in the subtracted frame. These residual values are treated as noise when data information is demodulated. Particularly, an edge of an object in a frame remains intense. Conversely, visual information is overlaid with data information of weak signal intensity $\alpha$ to avoid degradation of visual information, i.e., noise has large values, while data information has small values.
In the noise elimination method, the absolute value of each pixel within a cell of the subtracted frame is compared with a pre-defined threshold $\eta$. If the value is larger than the threshold $\eta$, this pixel is regarded to contain intense noise and is omitted for averaging the pixel values within the cell.

### 4 Experimental results

We experimentally evaluate BER performance of the expanded IS-VLC system. Table I lists the parameters of the components used in the experiment. We use a SHARP PN-Y475 as a display and a Raspberry Pi 3 Model B with Raspberry Pi Camera Module v2.1 as an image sensor. The experiment was carried out in a room at night, where fluorescent light was on. The distance between a display (Tx) and an image sensor (Rx) is 1.5 m. We use an 8-bit pixel value, that is, it can take from 0 to 255. A motion video named “sunflower” is used as visual information. In this video, a bee is flying over a sunflower, and the camera follows the bee. The center region of the frame abruptly changes because of the flying bee, while the other region gradually changes. The absolute category rating (ACR) [12] is used to evaluate the quality of visual information. After watching the visual information, evaluators of 13 males and 12 females scored it from 1 to 5. The quality is evaluated in terms of a mean opinion score (MOS).

Fig. 2 shows BER performance of the expanded and the conventional IS-VLC systems. The threshold $\eta$ is adjusted to reduce the BER. In Fig. 2(a), we can find that BER becomes low as the signal intensity becomes large. The expanded system can achieve better BER performance than the conventional system. In [13], a turbo code with a code rate of one-third can correct errors till the BER reaches $10^{-1}$. The performance of the expanded system satisfy the conditions of both a target BER (below $10^{-1}$) and a good MOS (over 4). For $\alpha = 5$, BER of the cell-position correction method becomes worse. This method reduces the region for averaging the pixel values for the large movement of the cell position. Therefore, it is preferable that the region of averaging the pixel values is enlarged when the signal intensity is large. Fig. 2(b) shows BER of each cell for $\alpha = 2$. The cell-position correction method can reduce BER in the peripheral region of the center. Hence, this method is effective for gradual change between successive frames. The noise

| Table I. Parameters used in the experiment. |
|------------------------------------------------|
| Frame rate of data information | 10 fps |
| Frame rate of visual information | 20 fps |
| Capture rate of an image sensor | 20 fps |
| Size of an data frame $(X, Y)$ | 1,900 pixels × 1,000 pixels |
| Cell construction $(M, N)$ | (19, 10) |
| Resolution of a display | 1,920 pixels × 1,080 pixels |
| Resolution of an image sensor | 1,920 pixels × 1,080 pixels |
| Distance between a display and an image sensor | 1.5 m |
| Signal intensity $\alpha$ | 1, 2, 3, 4, 5 |
| Visual information | Sunflower |
elimination method can improve BER in the whole region. This method is more robust for the movement between the successive frames.

5 Conclusions

In this letter, we expanded the IS-VLC system to support a motion video as visual information by introducing the cell-position correction and the noise elimination methods. From the experimental evaluation, we have clarified that these methods can improve BER performance for the motion video.

We evaluated the BER performance using only a motion video as visual information. However, it is not enough to evaluate the performance of the proposed system; therefore, in the future, we intend to use motion videos with various characteristics.

Acknowledgments

The authors would like to thank Prof. Yamazato of Nagoya University for his valuable suggestions. This work is partly supported by JSPS KAKENHI Grant Number 16K06348.