Correlation-based receiver for optical camera communications

Cristo Jurado-Verdu,[*] Vicente Matus, Jose Rabadan, Victor Guerra, and Rafael Perez-Jimenez

IDeTIC, Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain
[*]*cjurado@idetic.eu*

Abstract: In color-multiplexed optical camera communications (OCC) systems, data acquisition is restricted by the image processing algorithm capability for fast source recognition, region-of-interest (ROI) detection and tracking, packet synchronization within ROI, estimation of inter-channel interference and threshold computation. In this work, a novel modulation scheme for a practical RGB-LED-based OCC system is presented. The four above-described tasks are held simultaneously. Using confined spatial correlation of well-defined reference signals within the frame's color channels is possible to obtain a fully operating link with low computational complexity algorithms. Prior channel adaptation also grants a substantial increase in the attainable data rate, making the system more robust to interferences.

1. Introduction

Optical camera communications (OCC) have increased their relevance within the field of visible light communications (VLC) due to the widely-available cameras in a great number of devices (smartphones, tablets, surveillance systems, ...). Most of these cameras are based on rolling-shutter (RS) technique, which implements a row-by-row image scanning process [1]. Therefore, an LED light source, switching at a frequency higher than the shutter speed, will appear in the image as a series of dark and bright stripes, representing the binary data. Thus, this camera architecture provides to OCC a data rate much higher than its frame rate (fps) [2, 3]. OCC links need from several image processing stages to ensure a correct reception, such as detection and tracking of the light source (the region of interest, ROI), equalization of the optical power along the ROI, and threshold computation for demodulation. In ROI detection, blurring, dilating, and a variety of morphological operations have been proposed [4]. Other proposals [5] use anchor LEDs as spatial references for source detection, alignment, and spatial synchronization. However, this technique is not suitable for time synchronization in RS-based receivers.

For the non-uniformity of the received signal in the ROI, suggested solutions are based on searching the area in which this effect is partially mitigated [6], or by applying enhancing techniques to ameliorate it [7]. In the case of threshold computation, third order polynomial curve fitting (TOPF), iterative threshold (ITS), and quick adaptive scheme (QAS) have been evaluated in [8]. Furthermore, other techniques have been explored, such as entropy-based [9] and pixel-boundaries-based algorithms [10]. It is important to highlight that all the experiments above performed their threshold calculation frame by frame, with high computational resource consumption. Finally, the system has to spatially detect and synchronize the data packet within the ROI for correct data demodulation. In this case, two effects must be considered: the sampling frequency offset (SFO) of the camera, that varies slightly the sampling instants over time, and the blind period in which the signal cannot be acquired, either because the sensor is not being exposed (camera inter-frame processing), or because the signal does not fall within the ROI. Packet framing and repetition are used to address this issue [11]. Moreover, for RGB-based systems, the inter-channel cross-talk (ICCT), caused by the wavelength mismatch between the LEDs and the camera's Bayer filter, needs to be compensated as proposed in [12, 13], where a

training signal is employed.

In this paper, a correlation-based model is proposed for providing a unique solution to address all these issues. Using 2D-correlation processing over the images, the presented system is able to detect, track and identify the data source, even in case of spatial multiple access, with several sources included in the same image. Furthermore, using the correlation results, ICCT compensation and polynomial curve fitting for threshold computation can be obtained. Then, using the same correlation strategy, the system spatially synchronizes the data packet and provides the best-suited sampling spots for demodulation.

2. System description

The proposed OCC system is a broadcasting, RGB-based link with RS-camera as the receiver. The following subsections will describe in detail both transmitter and receiver structures.

2.1. Transmitter

Raw transmission data are partitioned on fixed-length packets. In order to keep data integrity, each packet is repeated during two frame times. Under this approach, data transmission is delimited in time within a super-frame structure. While the LED-lamp is sending data, a time window denoted as the Beacon-Only period, is reserved at the beginning of each super-frame for the transmission exclusively of N_b beacon packets. Additionally, when the emitter remains at an idle state, the beacons packets will be transmitted continuously. The functions of the beacon are: to ease the discovery stage, to uniquely identify each source (exploiting the spatial division multiplexing capability), and to help in the inter-channel cross-talk and threshold level computation. Figure 1(a) shows the beacon structure.

The length of the beacon, M, must be the same as the data packet. It consists of an N-length Hadamard sequence, corresponding to the number of wavelengths (R, G, and B in this case), followed by $M - N - 1$ single-color slots and a guard slot (depicted in black). This guard slot diminishes the effect of misidentifying two different sources. This configuration allows $N! \cdot (M - N - 1)^N$ different emitter identifiers within the image (where $N \geq (M - 1)$).

During the data transmission time, the packets being sent combine a synchronization signal that occupies exclusively one of the independent channels, and the OOK-modulated payload distributed over the other ones. This synchronization signal marks the start, end of the packet, and the optimal sampling instants (Fig. 1(b)).

2.2. Rolling shutter-based receiver

The receiving process can be split into three consecutive stages: discovery, training, and acquisition. The receiver will use two 2D-correlation templates: the beacon, and the synchronization templates, derived from the corresponding transmitter's signals (Fig. 1). They have a fixed width of w pixels, and a stripe height, $h = t_{chip}/t_{row}$. Where t_{chip} and t_{row} are the symbol and sensor's row delay times respectively. This h optimizes the correlation output, and increase the resilience to the broadening effect that the overlapping exposure of consecutive rows could generate in the stripe height. The total beacon height (in pixels), $H_{beacon} = h \cdot M$, must be less or equal to the half of the expected vertical height of the lamp's projection, H_{lamp}, within the frame, $H_{lamp}/2 \geq H_{beacon}$. The computation of H_{lamp} is discussed in [14]. This constraint ensures that a beacon could be recovered within the lamp's projection after the processing of two frames, and limits the final throughput, R_b (Eq. (1)).

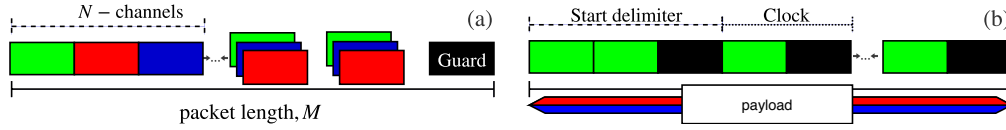$$Rb \leq \frac{H_{lamp}}{2} \cdot \frac{1}{h} \cdot \frac{fps}{2} \cdot N \tag{1}$$

Fig. 1. Transmitter signals. (a) Beacon signal. (b) Data signal.

2.2.1. Discovery stage

In this stage, the receiver correlates a detection template that groups two beacon templates. If the maximum Pearson correlation coefficient, ρ, (Eq. (2)) exceeds the imposed detection threshold value, ρ_{th}, a source is considered as successfully detected.

$$\rho(x, y) = \frac{\sum_{x',y'}(T'(x', y') \cdot I'(x + x', y + y'))}{\sqrt{\sum_{x',y'} T'(x', y')^2 \cdot \sum_{x',y'} I'(x + x', y + y')^2}} \tag{2}$$

T and I are the template and image frames respectively, while x and y are the pixel coordinates. After detection, the receiver proceeds to the training stage.

2.2.2. Training stage

In this stage, the receiver locates the beacon template within the cropped detected ROI. Taking advantage of the beacon structure (independently switched channels), RGB to Bayer gains can be directly obtained from channel samples of N frame pictures.
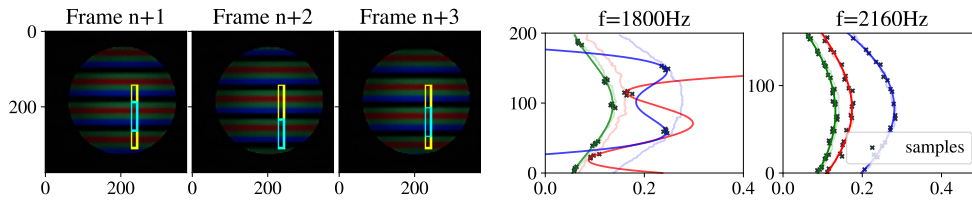

Fig. 2. (a) Inter-frame motion example. (b) Comparison of training samples (dark dots) and curve fitting for motionless $F_{tx} = 1800$ Hz and $F_{tx} = 2160$ Hz.

An example of the capture of 3 frames used for training is shown in Fig. 2(a). The yellow rectangle highlights the cropped ROI area, while the beacon is framed with cyan borders. As can be appreciated, the beacon is not fixed among frames but advances or falls behind its previous position. As has been stated in RS systems, there is an implicit relative motion of just a few rows per a large number of frames caused by the camera's SFO, v_{sfo}. This motion has been generally considered as an issue. Nevertheless, the proposed system benefits from this phenomenon since it allows to increase the spatial dispersion of the training samples between frames. Furthermore, to speed up this inter-frame motion a $v_{design} = [N_s \bmod M] \cdot h$ (pixels/frame), can be forced by selecting the length of the beacon template accordingly over the total number of available stripes in the frame, N_s. Taking this parameter into consideration, the total inter-frame motion is, $v_{inter} = v_{design} + v_{sfo}$.

After processing L frames, the system performs a third order polynomial fitting to obtain the cross-talk matrix and the threshold for the entire ROI and proceeds to the next stage. Figure 2(b) depicts the collected samples from 5 consecutive frames and the output fitted curve when, v_{design} was set to zero (left graph). In that case, it can be observed that samples tend to form local groups (dark dots) instead of evenly distribute over the entire ROI, as it occurs when motion is forced (right graph). If a beacon template is not found, the receiver will restart.

2.2.3.   Acquisition stage

In this stage, the receiver correlates both the beacon and the synchronization templates with the ROI. When a sync template is found, the receiver performs the enhancing exclusively over that area. This reduces the computational load needed for pre-processing the entire image. Finally, the binarization process and data assembly are performed. If any of the templates are not found during this stage, the receiver will restart the discovery stage.

3.   Experiment and results

Fig. 3. Experiment setup.

As a proof of concept, a testbed was put together to measure the system performance as shown in Fig. 3. The transmitter is an RGB-LED lamp driven by an ST NUCLEO board. The signal was recorded using a Logitech C920 webcam. The videos were processed off-line using OpenCV libraries. In order to evaluate the operation of this proposal, it was carried out a series of experiments for different distances, d, and frequencies, f. The camera fps was set to 30 with full HD resolution (1080p). The exposure time was set to the minimum available, $300\,\mu s$. The measured row delay time was $31.4\,\mu s$. The camera's vertical Field of View (VFOV) was 43.3°. The ISO was set to 100, and the white balance correction to 6500K. The template's column width, w, was set to 15 pixels. Finally, the selected frequencies: 2160, 3240, 1800 and 2700 Hz (with stripes' height of 15, 10, 18 and 12 pixels respectively) can be grouped into a forced (first two) and non-forced motion sets. The physical height of the light source was 8 cm. For each pair (d, f), two different video recordings were made (10 minutes) while the source was transmitting beacons continuously and when it was sending pseudo-random packets of 10 bits. The last recording was performed three times, resulting in 1.25×10^5 received raw bits per experiment.
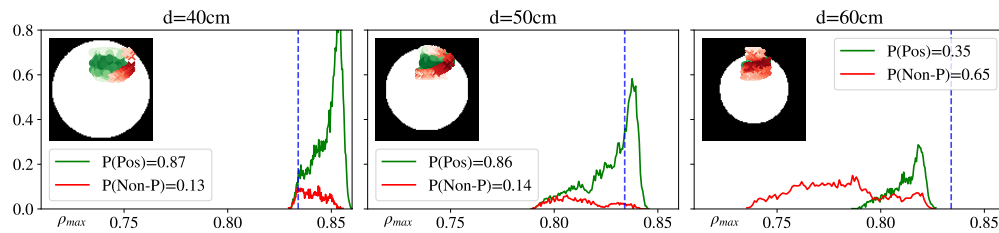

Fig. 4. Maximum Pearson correlation coefficient, $\rho_{\max}$. $F_{\text{tx}} = 2160\,\text{Hz}$.

The system's precision detecting legitimate sources with a certain degree of confidence was evaluated (positive case). To perform this evaluation, the maximum Pearson correlation coefficient between each frame and the template was obtained. Then, it was classified into a positive or non-positive sample collection. Samples are considered positive when the beacon template completely fits within the source's projection. Figure 4 depicts the detected position associated to the maximum correlation (green for the positive cases), and the histogram of the correlation coefficient for both positive and non-positive cases weighed by their corresponding *a priori* probabilities. It can be highlighted that as distance increases, both the source's projected

area and the number of positive samples diminish. Moreover, the average correlation value also decreases. This occurs because the template is forced to be detected at the lamp's center, where it can fit entirely. Nonetheless, at this position, the pulse broadening effect is higher, leading to a lower correlation output value.

Based on these samples, a detection threshold is selected. Lowering the threshold level will increase the number of false positives (detecting a non-positive sample as a source). Thus, the receiver's precision will drop. Otherwise, if the threshold raises, the miss rate would rapidly grow. This has implications on the average source detection time that can be expressed in terms of the number of frames needed for source detection, $N_{\text{detection}}$. Figure 5(a) plots the precision/recall
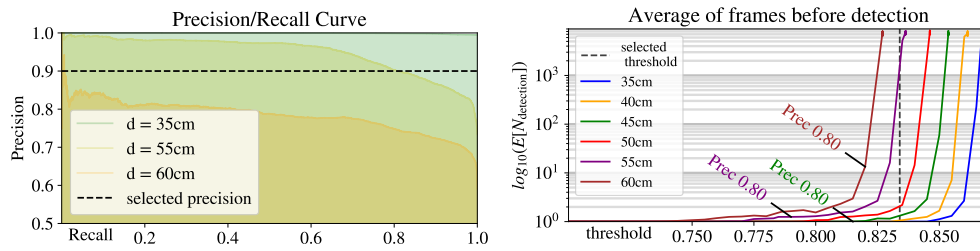

Fig. 5. (a) Precision-Recall curve. (b) Average of $N_{\text{detection}}$. $F_{\text{tx}} = 2160\,\text{Hz}$.

curve of the system. The dashed black line sets the minimum precision set as the design criteria of 0.9 (90 percent of the detected sources will belong to the true positive case). Figure 5(b) plots the average number of frames prior detection, $E[N_{\text{detection}}]$, against the detection threshold. In the extreme case in which the ROI height is comparable to the source projection over the image, there would be a higher probability of missing the detection. If the inter-frame motion were low, the probability of detecting the source on the next frame would also be small. Thus, there is an inverse relationship between the inter-frame motion and the variance of $N_{\text{detection}}$ for those cases. For instance, if the receiver captures a beacon halfway through its transmission, it will have to wait a long time for detection due to the scarce inter-frame motion.
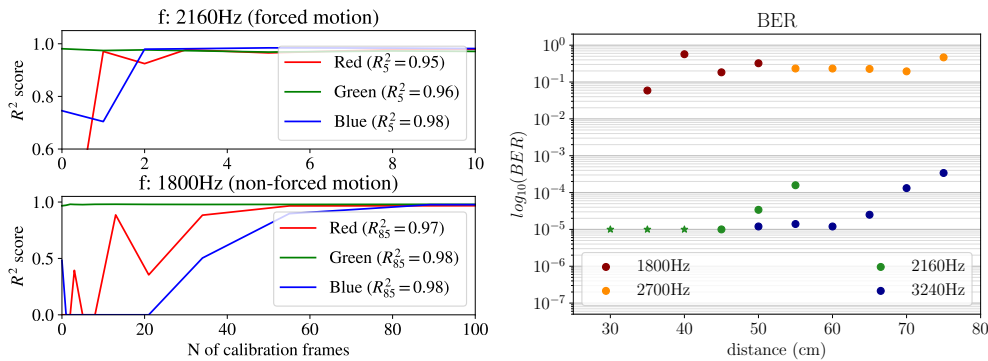

Fig. 6. (a) R^2 determination coefficient versus number of frames for training. (b) Bit error rate versus distance (star points indicate that no errors were detected under the experiment conditions).

Then, to evaluate the performance of the training, the R^2 determination coefficient is used. For both frequency sets, a set of samples were collected through N frames, fitted and compared with independently captured images from each channel. Figure 2(b) represents the third polynomial fit against the real curve (lighter lines), obtained with independent image captures. As can be

seen in Fig. 6(a), the non-forced motion frequency set needs more frames for training due to sampling clustering, to obtain an optimal R^2 determination coefficient of the fitting.

Finally, Bit Error Rate (BER) performance is presented in Fig. 6(b). It was evaluated using 0.834 as the detection threshold (system precision of 0.9), and four calibration frames. It can be seen that the BER decrements with the distance. As has been mentioned, this is related to the rising of the stripe broadening effect, which has a harmful effect on calibration, since there is a higher probability to obtain color-mixed samples.

4. Conclusions

In this work, an experimental evaluation of an RGB LED-based OCC system is presented. It uses the green channel for data synchronization, while the red and blue ones carry OOK-modulated data. Furthermore, a beacon-based detection scheme is proposed and evaluated. The processing algorithms for ROI detection, source identification, training, and packet synchronization are combined into a single correlation-based procedure. This technique finds the best ROI in terms of the least pulse broadening (inter-symbol overlapping), improving the BER performance. It also carries out the ICCT mitigation and enhancing, only within the data region, reducing the computational load. Experimental results demonstrated that the proposed system is able to achieve 300 bps (over a transmission span up to 0.7 m, with a constant BER lower than 1×10^{-4}. However, higher data rates or higher distances could be achieved by increasing the physical size of the lamp and the framing structure or using spatially-multiplexed sources.

Funding

References

1. H. Aoyama and M. Oshima, "Line scan sampling for visible light communication: Theory and practice," in *IEEE International Conference on Communications (ICC),* (IEEE, 2015), pp. 5060–5065.
2. N. T. Le, M. A. Hossain, and Y. M. Jang, "A survey of design and implementation for optical camera communication," Signal Process. Image Commun. **53**, 95–109 (2017).
3. Z. Ghassemlooy, P. Luo, and S. Zvanovec, "Optical Camera Communications," in *Optical Wireless Communications: An Emerging Technology,* M. Uysal, C. Capsoni, Z. Ghassemlooy, A. Boucouvalas, and E. Udvary, eds. (Springer, 2016), pp. 547–568.
4. J.-W. Lee, S.-J. Kim, and S.-K. Han, "Multi-Level Optical Signal Reception by Blur Curved Approximation for Optical Camera Communication," in *2017 Opto-Electronics and Communications Conference (OECC) and Photonics Global Conference (PGC),* (2017).
5. A. D. Griffiths, J. Herrnsdorf, M. J. Strain, and M. D. Dawson, "Scalable visible light communications with a micro-LED array projector and high-speed smartphone camera," Opt. Express **27**, 15585 (2019).
6. J. He, Z. Jiang, J. Shi, Y. Zhou, and J. He, "A Novel Column Matrix Selection Scheme for VLC System With Mobile Phone Camera," IEEE Photonics Technol. Lett. **31**, 149–152 (2019).
7. W. Guan, Y. Wu, C. Xie, L. Fang, X. Liu, and Y. Chen, "Performance analysis and enhancement for visible light communication using CMOS sensors," Opt. Commun. **410**, 531–551 (2018).
8. K. Liang, C.-W. Hsu, C.-W. Chow, C.-Y. Chen, Y. Liu, S.-H. Chen, and H.-Y. Chen, "Comparison of thresholding schemes for visible light communication using mobile-phone image sensor," Opt. Express **24**, 1973 (2016).
9. K. Liang, C.-W. Chow, Y. Liu, and C.-H. Yeh, "Thresholding schemes for visible light communications with CMOS camera using entropy-based algorithms," Opt. Express **24**, 25641–25646 (2016).
10. Z. Zhang, T. Zhang, J. Zhou, Y. Lu, and Y. Qiao, "Thresholding Scheme Based on Boundary Pixels of Stripes for Visible Light Communication With Mobile-Phone Camera," IEEE Access **6**, 53053–53061 (2018).
11. T. Nguyen, C. H. Hong, N. T. Le, and Y. M. Jang, "High-speed asynchronous Optical Camera Communication using LED and rolling shutter camera," in *International Conference on Ubiquitous and Future Networks, ICUFN,* (2015).
12. K. Liang, C.-W. Chow, and Y. Liu, "RGB visible light communication using mobile-phone camera and multi-input multi-output," Opt. Express **24**, 9383–9388 (2016).
13. P. Luo, M. Zhang, Z. Ghassemlooy, H. Le Minh, H. M. Tsai, X. Tang, L. C. Png, and D. Han, "Experimental Demonstration of RGB LED-Based Optical Camera Communications," IEEE Photonics J. **7**, 1–12 (2015).
14. P. Chavez-Burbano, V. Guerra, J. Rabadan, D. Rodriguez-Esparragon, and R. Perez-Jimenez, "Experimental characterization of close-emitter interference in an optical camera communication system," Sensors **17**, 1561 (2017).