A Tensor-Based Optical Camera Communication (OCC) System With Joint Data Detection and Video Restoration

Leandro R. Ximenes, Member, IEEE, and Bruno de. A. Larêdo

Abstract

The Optical Camera Communication (OCC) technology allows Visible Light Communications (VLC), one of the key technologies for the Sixth Generation (6G) of mobile communications, to be used with image sensors on the receiver side. For Screen-to-Camera (S2C) applications, in which Light-Emitting Diode (LED) screens are used to send information symbols, increasing either (transmitted) image resolution or video frame rate may negatively impact the bit error performance and the visual quality of the OCC-encoded videos. This work proposes a completely new OCC/S2C transmission and reception system, in which it is possible to recover both the transmitted symbols and to restore the video quality at the same time. For this dual-task, a Parallel Factor Analysis (PARAFAC) decomposition is applied to model an OCC system for the first time in the literature. Computational simulations corroborate the correct formulation of the proposed models, propositions and algorithms, in addition to demonstrating that increasing the number of frames of the OCC-encoded video can generate diversity gains. This work also includes discussion on the practical aspects of data rate, data volume, and computational complexity of the receiver, accordingly to the system’s parameters.
Index Terms

Optical communication, signal processing, camera, video signal processing

I. INTRODUCTION

Optical Camera Communication (OCC) is a recent technology with a high potential for rapid implementation in society. This technology was standardized in [1] as a substandard of the Visible Light Communication (VLC) technology. What differentiates OCC from VLC is that in the first one, the receiver is not based on discrete photodiodes (PDs), but on Image Sensors (IS), typically found in all modern smartphones [2]. Since there is a dizzying growth of the market penetration of these devices for the following years, the OCC technology is highly attractive and with several applications already foreseen or tested, such as vehicular communications [3], [4], indoor location positioning/localization [5], [6], broadcasting [7], digital signage, among others [8]. OCC, and VLC as a whole, are considered potential technologies of the Sixth Generation (6G) of wireless communications [9], [10].

OCC systems can generally be classified according to the number of elements at each end of the communication link. On the source side, the Screen-to-Camera (S2C) category, also called display-camera [11], [12] or simply screen-camera [13], predicts using hundreds or thousands of LEDs as transmitters [8]. This technology thus has tremendous application potential since the number of commercial LED screens has grown even faster than the number of smartphones, driven by the market growth of these devices, digital billboards, and modern display-based consumer electronics [14]. Image sensors are usually based on Complementary Metal-Oxide-Semiconductor (CMOS) or Charge-Coupled Device (CCD) technologies on the receiver side.

By employing LEDs and image sensors, OCC/S2C systems provide an exciting integration of wireless communication and image processing. The communication design needs to ensure modulation and coding schemes that guarantee the desired bit error rate, data rate, and other link
requirements. For the parallel task of video transmission, one must encode the binary message into the video pixels without impairing the subjective visual quality perceived by the human eye.

Due to the limitation of the camera’s frame rate, in the order of less than a hundred frames per second for low-cost smartphones, the concept of Optical Multiple-Input Multiple-Output (MIMO) is usually exploited to enhance data rate. Thanks to the high number of LEDs on a screen and the high number of photodetectors present in an image sensor, ultra-massive MIMO theoretical channel capacity can be achieved with S2C systems [13].

In spite of such promising benefits, the duality in the transmission of symbols and visible image conditions the performance of an S2C system to two distinct but related propagation issues. The first one concerns the so-called interpixel interference (IPI), which can generally be summarized as the case in which the light intensity of a pixel hinders the identification of the intensity of its neighbors. In addition, in general, IPI can be related both to the blooming effect and to classical image degradation phenomena, such as optical blurring or motion blurring. For the blooming effect, the luminous power of a given pixel is strong enough to illuminate the neighboring captured pixels. The works [15], [16] propose ways to mitigate this issue. In this case, even classic image restoration methods can be used to mitigate the blurring effects, generated either by the low resolution or by the low capture rate of the optical camera.

The second propagation issue is the negative visual impact of embedding information symbols on the pixels of encoded video. Unlike the previous issue, inherently caused by hardware limitations or due to non-idealities of the optical channel, this issue is exclusively related to the OCC coding and modulation.

Several modulation schemes, coding techniques, and receivers for OCC have already been proposed in the literature ([17], [18] and works therein). Digital modulations such as On-Off Keying (OOK), Pulse Position Modulation (PPM), Quadrature Amplitude Modulation (QAM), and Color-Shift Keying (CSK) [1] are implemented [19], [20], as well as different sampling methods such as Screen-Nyquist sampling and Rolling-Shutter Nyquist sampling [1].
All these modulations specifically proposed for VLC or OCC systems inherently carry characteristics that affect the perception of light by human vision to a greater or lesser degree. For example, the CSK modulation prevents flickering by transmitting at constant power. In contrast, the UFSOOK modulation [21] cleverly handles the need for a high pulse rate in the transmission (to avoid flickering perception) with an undersampling approach at the receiver (to deal with low capture rates of the commercial cameras). While the accepted threshold of screen flickering is less critical in S2C systems [18], their satisfactory performance inevitably ends up being also defined by the deterioration of the subjective visual quality caused by the symbol-pixel modulation [22].

This work proposes a new S2C/OCC system based on tensor modeling. More precisely, this is the first time the OCC (or S2C) technology has been combined with a tensor modeling. This tensor approach allows the concomitant processes of symbol detection and video restoration on the video captured by the optical camera. In other words, it proposes a system with a better compromise between data estimation performance and video transmission quality.

The use of tensor models have seen considerable growth in the past decades, ranging from applications in wireless communications [23], [24], signal processing and machine learning [25], Big Data analysis [26], [27], biology and health studies [28], [29], and many other fields.

Perhaps the most significant advantage of using tensors in system modeling lies in relaxing parameter estimation conditions. For example, in wireless communications, tensor modeling of the transmission system usually allows the receiver to estimate symbol and channel matrices from multidimensional signals under less stringent conditions than in purely matrix-based approaches.

Several tensor models have been proposed in the literature. Although the vast majority have been proposed in the last two decades ( [25]), perhaps the best-known decomposition is the PARAFAC decomposition [30]. In the trilinear PARAFAC decomposition, up to three matrix factors can be jointly estimated, whose estimation solution is unique under certain relaxed conditions.
Briefly, the combination of tensor modeling in S2C/OCC systems is responsible for the following contributions of this article:

- A complete OCC/S2C communication system based on the PARAFAC decomposition is developed. In this work, information signals modulate each LED/pixel of a digital screen using a Khatri-Rao Space-Time (KRST) coding [31]. More precisely, arbitrary digital video frames, such as from a movie, are encoded in blocks by sequences of symbols to generate a video tensor. The tensor’s multidimensionality comes from the original video’s pixel resolution, the video frames, and the symbol time-slots. System parameters, as the number of video frames, define the trade-off between Symbol Error Rate (SER), video restoration performance, and receiver’s complexity.

- A non-iterative blind receiver named OCC-KRF is proposed for this scheme. Unlike other OCC receivers in the literature, OCC-KRF has a special dual function: it allows at the same time the estimation of information symbols and the image restoration of the captured video.

- Formulas for data rate, generated data volume, and computational cost are derived. Design rules are proposed to assure the essential uniqueness and identifiability conditions of the proposed system. It is also shown that the receiver easily avoids all ambiguities of the PARAFAC decomposition if it knows at least a single (pilot) frame of the original video.

This work is segmented into eight sections. In Section II, the definitions of mathematical notations used in this work are presented, along with a brief review of the PARAFAC tensor decomposition. In Section III, the proposed system is developed. Section IV describes how the OOK, PPM, and PWM modulations can be compatible with the proposed scheme are presented. In Section V, the OCC-KRF algorithm is introduced, along with data rate, data volume, and receiver complexity calculations. In Section VI, numerical analysis through computer simulations is performed to solidify the contributions’ validity. After the conclusion in Section VII, the Appendix brings three design rules in the form of mathematical propositions dedicated to
establishing uniqueness and identifiability conditions for the models and systems of this work.

II. NOTATIONS AND REVIEW ON PARAFAC

Scalars, column vectors, matrices, and tensors are denoted by lower-case (x), boldface lower-case (x), boldface capital (X), and calligraphic (X) letters, respectively. X^T, X* and X^† are the transpose, the conjugate and the pseudoinverse of X, while \( \hat{X} \) denotes an eventual estimate. The \( i \)-th row-vector of X is given by \( X_i \cdot \), while the \( j \)-th column of is given by \( X \cdot j \). Moreover, \( 1_{I \times J} \) defines the \( I \times J \) all-ones matrix, and \( I_J \) is the identity matrix of \( J \times J \) size.

The Khatri-Rao product K of two arbitrary matrices \( A^{(1)} \in \mathbb{C}^{I_1 \times R} \) and \( A^{(2)} \in \mathbb{C}^{I_2 \times R} \) is defined as

\[
K \triangleq A^{(1)} \odot A^{(2)} \in \mathbb{C}^{I_2 I_1 \times R},
\]

\[
= \left[ \begin{array}{cccc}
A^{(1)}_1 \otimes A^{(2)}_1 & A^{(1)}_2 \otimes A^{(2)}_2 & \cdots & A^{(1)}_R \otimes A^{(2)}_R
\end{array} \right],
\]

where \( \otimes \) is the Kronecker product. For a \( I_1 I_2 \) concatenated array index, its element ordering follows the lexicographic order \((i_2 - 1)I_1 + i_1 = 1, \cdots, I_1 I_2\), with \( i_1 = 1, 2, \cdots, I_1 \) and \( i_2 = 1, 2, \cdots, I_2 \).

For a third-order tensor \( X \in \mathbb{C}^{I \times J \times K} \) with a PARAFAC [30] decomposition \( X = [A^{(1)}, A^{(2)}, A^{(3)}; R] \), its mode-1, mode-2 and mode-3 unfoldings are respectively given by

\[
X^{(1)} = (A^{(3)} \odot A^{(2)}) (A^{(1)})^T,
\]

\[
X^{(2)} = (A^{(1)} \odot A^{(3)}) (A^{(2)})^T,
\]

\[
X^{(3)} = (A^{(2)} \odot A^{(1)}) (A^{(3)})^T,
\]

where \( R \) is the tensor-rank and the number of columns of \( A^{(1)} \in \mathbb{C}^{I_1 \times R} \), \( A^{(2)} \in \mathbb{C}^{I_2 \times R} \), and \( A^{(3)} \in \mathbb{C}^{I_3 \times R} \).

The powerful essential uniqueness property of the PARAFAC decomposition is summarized in Theorem 1. Notations and definitions used in the following theorems follow those defined in the previous paragraphs of this section.
Theorem 1 (c.f. [30]). Given the model $\mathcal{X} = \{A^{(1)}, A^{(2)}, A^{(3)}; R\}$, then the property of the essential uniqueness of the PARAFAC decomposition means that its matrix factors are unique up to column scaling and permutation ambiguities. In other words, there is a permutation matrix $\Pi$ and diagonal scaling matrices $\Lambda_A$, $\Lambda_B$, and $\Lambda_C$ such that

$$\mathcal{X} = \{\hat{A}^{(1)}, \hat{A}^{(2)}, \hat{A}^{(3)}; R\},$$

where

$$\hat{A}^{(1)} = A^{(1)} \Pi \Lambda_A, \quad \hat{A}^{(2)} = A^{(2)} \Pi \Lambda_B, \quad \hat{A}^{(3)} = A^{(3)} \Pi \Lambda_C,$$

$$\Lambda_A \Lambda_B \Lambda_C = I_R.$$ (8)

Sufficient uniqueness conditions are reviewed in the Theorems 2, 3, and 4.

Theorem 2 (c.f. [32]). Given $\mathcal{X} = \{A^{(1)}, A^{(2)}, A^{(3)}; R\}$, then this decomposition is essentially unique if $k_{A^{(1)}} + k_{A^{(2)}} + k_{A^{(3)}} \geq 2R + 2$, where $k_{A^{(i)}}$ denotes the Kruskal rank of $A^{(i)}$.

Theorem 3 (c.f. [33], [34]). Given $\mathcal{X} = \{A^{(1)}, A^{(2)}, A^{(3)}; R\}$, let $I_1 \geq R$. Then this decomposition is essentially unique, almost surely, if $I_2(I_2 - 1)I_3(I_3 - 1) \geq 2R(R - 1)$.

Theorem 4 (c.f. [35]–[37]). Given $\mathcal{X} = \{A^{(1)}, A^{(2)}, A^{(3)}; R\}$, let $I_1 \geq R$ and $\min(I_2, I_3) \geq 3$. Then this decomposition is essentially unique, almost surely, if $(I_2 - 1)(I_3 - 1) \geq R$.

III. PROPOSED TENSOR-BASED OCC SYSTEM

In our proposed OCC/S2C system, each pixel of a digital video of resolution $J$ pixels by $L$ pixels is modulated by a string of $S$ symbols. Each symbol is formed by $P$ samples (pulses), such that $K = SP$ is the number of samples per pixel. In other words, real baseband symbols stored in the matrix $S \in \mathbb{R}^{K \times JL}$ determine the information to be encoded into $JL$ pixels.
Let this video comprise a set of $F$ frames, and that $x_f \in \mathbb{R}^{JL \times 1}$ is a vectorized form of the $f$-th frame. Video matrix $X = [x_1 \ x_2 \ \cdots \ x_F] \in \mathbb{R}^{F \times JL}$ is the rearranged block of digital video in which it is desired to encode the message contained in $S$.

The OCC-encoding proposed here is based on the KRST coding [31], so we have the following coded video for displaying

$$\tilde{X} = (S \ast X) \in \mathbb{R}^{KF \times JL}. \quad (9)$$

Realize that thanks to the Khatri-Rao operation in (9), the $F$ frames of $X$ are replicated by an $K$ factor. In other words, the native encoded video becomes $K$ lengthier than the original one, and then must be played at a speed $K$ times greater.

Consider that during transmission, the video $\tilde{X}$ is degraded accordingly to the degradation model

$$Y = H\tilde{X}^T, \quad (10)$$

where $Y$ is the captured video. The degradation matrix $H$ may be given by

$$H = DWB \in \mathbb{R}^{MN \times JL}, \quad (11)$$

where $D \in \mathbb{R}^{MN \times JL}$, $W \in \mathbb{R}^{JL \times JL}$ and $B \in \mathbb{R}^{JL \times JL}$ represent respectively the generic decimation, warping and (optical and motion) blurring matrices [38]. Note that $H$ introduces decimation if $MN < JL$, where $M \times N$ is the spatial resolution of the received video.

Combining (10) and (11), the noiseless video $Y$ generated by the Image Sensor (IS) is given by

$$Y = H(S \ast X)^T \in \mathbb{R}^{MN \times KF}, \quad (12)$$

where the rows of $Y$ represent the spatial domain (i.e., pixels) and the columns the time domain (i.e., video frames). In practice, the acquisition of $Y$ is indeed done by capturing $KF$ video frames of resolution $M$ by $N$ pixels.
Comparing (12) with the transpose of (3), and making the correspondence \((X^{(1)}, A^{(1)}, A^{(2)}, A^{(3)}) \leftrightarrow (Y^T, H, X, S)\), we can define a third-order PARAFAC model \(\mathcal{Y} \in \mathbb{R}^{MN \times F \times K}\) as

\[
\mathcal{Y} = [H, X, S, JL].
\]  

(13)

The concept of constructing \(\mathcal{Y}\) is illustrated in Fig. 1.

Assuming the presence of additive noise and interference in the capturing process, the general system model is given by

\[
\tilde{Y} = Y + N,
\]

(14)

where \(N \in \mathbb{R}^{MN \times K \times F}\) is the additive noise tensor, which may be thermal, shot noise, sampling noises, among others [17].

From (13) the mode-1, mode-2, and mode-3 unfoldings of \(\mathcal{Y}\) can be given respectively by

\[
Y^{(1)} = (S \circ X) H^T \in \mathbb{C}^{FK \times MN},
\]

(15)

\[
Y^{(2)} = (X \circ H) S^T \in \mathbb{C}^{MNF \times K},
\]

(16)

\[
Y^{(3)} = (H \circ S) X^T \in \mathbb{C}^{KMN \times F}.
\]

(17)

The proposed OCC system is summarized in Fig. 2.
IV. PULSE MODULATIONS

Our proposed OCC transmission scheme is compatible with different pulse modulations, such as the different variants of Pulse Position Modulation (PPM), Pulse Amplitude Modulation (PAM), and Pulse Width Modulation (PWM), since the choice of modulation changes only the symbol matrix. The discussion of the OOK, PWM, and PPM modulations within the proposed scheme is presented in this section.

The signal patterns of the Undesampling Frequency Shift On-Off Keying (UFSOOK) [21] and Undesampling Phase Shift On-Off Keying (UFSOOK) [39] modulations could also be emulated with the structure of $S$. For these modulations, the $P$ parameter can be used as the upsampling factor of the transmitted signals.
A. *On-Off Keying - OOK*

The simplest of the modulations used for light communication is the On-Off Keying (OOK). In the OOK modulation two LED states represent the two possible symbols. That is, Bit 1 can represent the non-null pulse and Bit 0 the null pulse. Fig. 3 illustrates the OOK modulation for a symbol of \( P \) pulses.

![Figure 3: OOK modulation](image)

For OCC systems, uncoded OOK modulation is typically undesirable, as a long string of zeros can mislead the human viewers to think of dead pixels. Appropriate compensation symbols or coding schemes can be used to control LED dimming [17], [40].

B. *Pulse Width Modulation - PWM*

In PWM modulation, the symbol determines the duration of the transmission pulses at high or low levels. The ratio of the time interval of the high-level pulse (non-zero pulse) to the symbol period is called Duty Cycle. The Duty Cycle (D), according to Fig. 4, is given by:

\[
D = \frac{P_H}{P} \times 100\%,
\]  

(18)
where $P$ is a parameter known to the receiver, and $P_H$ is based on the symbol. For example, for $P = 4$, four possible symbols \{$S_0, S_1, S_2, S_3$\} can be generated using $P_H = \{P/4, P/2, 3P/4, P\}$, which would generate duty cycles of 25\%, 50\%, 75\% and 100\% respectively.

![Figure 4: PWM modulation](image)

**C. Pulse Position Modulation - PPM**

In binary Pulse Position modulation (2-PPM), the two possible symbols within the alphabet are defined by the pulses’ position in the symbol interval.

In the PPM modulation of Fig. 5, the position of the pulse at low-level is determined by the bit to be transmitted, while all the remaining pulses are at high-level. Here low level means that pixel is off, while high level means the pixel is on. For Bit 0, the low-level pulse occurs at the beginning of the symbol, while for Bit 1, the low-level pulse occurs at the last pulse.

As anyone might suspect, higher-order PPM can easily be implemented with few adjustments.

In the PPM, it is not possible to control the LED dimming, as the level of direct current (DC) of the signal is the same for any of the transmitted symbols. This characteristic is less critical for S2C systems than for conventional VLC, as ambient lighting is not the main goal.
V. OCC-KRF RECEIVER

In possession of $\hat{Y}$, a receiver can estimate the symbols in $S$ while restoring the video $X$. This dual-task is achieved by the OCC Khatri-Rao Factorization (OCC-KRF) algorithm, described in Alg. 1.

The OCC-KRF receiver is based on a non-iterative framework, where degradation matrix $H$ in (11) must be known a priori. The OCC-KRF algorithm it is based on the Least Squares Khatri-Rao Factorization (LSKRF) algorithm adapted from [23], [24].

Given that $H$ is known, then the Khatri-Rao product $S \diamond X$ in (15) can be estimated by a Least Squares (LS) method (Step 2 in Alg. 1). The OCC-KRF algorithm then decompose this product into its two matrix factors (Step 3 in Alg. 1).

In the output of OCC-KRF, the estimated matrices $\hat{S}$ and $\hat{X}$ may have scaling ambiguities on their columns, as by Theorem 1. Let the correspondence $(A^{(1)}, A^{(2)}, A^{(3)}) \leftrightarrow (H, X, S)$ still be valid. Once $H$ is known, then $H = \hat{H}$, which means from (7) that $A_A = \Pi = I_R$. 

Figure 5: 2-PPM modulation
Therefore, there is only an arbitrary diagonal matrix $\Delta \in \mathbb{C}^{JL \times JL}$ that satisfies (8), i.e.,

$$X = \hat{X}\Delta, \quad S = \hat{S}\Delta^{-1}. \tag{19}$$

Ambiguity $\Delta$ is easily avoided if $X_f$ is known beforehand by the receiver, because $X_f = \hat{X}_f\Delta$. The ambiguity removal is done Steps 4 and 5 of Alg. 1. Note in (21) that $\hat{X}_f$ (and $X_f$) must not have any zero element.

**A. Considerations on Frame Rate, Data Volume, and Symbol Rate**

In S2C systems, information symbols are encoded into video pixels and then recovered by an optical camera receiver, a priori without impairing perceived visual quality if the screen’s refresh rate is greater than 70 Hz [41].

Let $r_S$ be the video effective playback rate, $r_V$ the original (native) uncoded video frame rate, and $r_C$ the capture rate of the IS installed at the receiver. Due to the KRST encoding, $r_S$ becomes $K$ times greater than $r_V$, i.e., the video must be played at $r_S = K r_V$. On the other hand, by definition, $r_V = F/T$ frames per second, where $T$ is the video block duration (in seconds). Finally, considering the acquisition by the IS, one can establish that

$$r_C \geq r_S = K F/T. \tag{23}$$

For $r_C > r_S$, the video $Y$ is oversampled at the receiver.

For an IS resolution of $M \times N$ pixels, with $b$ bits of quantization, the total volume of data (in Bytes per second) to be processed by the receiver is

$$V_T = (b/8) MN r_c \geq (b/8) MNKF/T. \tag{24}$$

Regarding the symbol rate $r$, since $JL$ screen’s pixels carry $S$ symbols each per $T$ seconds, then $r$ is equal to

$$r = JLS/T = JLSr_V/F = JLr_S/(FP). \tag{25}$$
Algorithm 1 OCC-KRF

1: Organize $\tilde{Y}$ into its mode-1 unfolding $\tilde{Y}^{(1)}$;

2: From (12), do the LS minimization: $\hat{Y} = \tilde{Y}^{(1)}(HT)^\dagger \approx S \odot X \in \mathbb{C}^{JL \times KF}$

3: For $i = 1, \cdots, JL$:
   
   i. Take the $i$-th row of $\hat{Y}$, and rearrange its $KF$ entries to form the rank-one matrix $\hat{Y}_i \in \mathbb{C}^{K \times F}$, such that its $(k, f)$ element, for $\{k, f\} \geq 1$, corresponds to the $(k - 1)F + f$ element of the original row vector.

   ii. Compute the Single Value Decomposition (SVD) of $\hat{Y}_i$, i.e., $\hat{Y}_i = UDV^T \approx \hat{S}^T \hat{X} \in \mathbb{C}^{K \times F}$.

   iii. Find the $i$-th rows of $\hat{S}$ and $\hat{X}$ from the rank-one approximation:

   $$\hat{S}_i = U_{:1}^T, \quad \hat{X}_i = V_{:1}^T,$$

   where $U_{:1}$ and $V_{:1}$ are respectively the first columns of $U$ and $V$. Neglect the singular values in $D$.

4: For an arbitrarily $f$-th known video frame, find the ambiguity matrix

$$\Delta \approx (X_f) ./ (\hat{X}_f),$$

where operator $./$ indicates the element-wise division.

5: Remove column ambiguities on $\hat{S}$ and $\hat{X}$:

$$\hat{X} \leftarrow \hat{X} \Delta^{-1}, \quad \hat{S} \leftarrow \hat{S} \Delta.$$  

In other words, comparing (24) to (25), increasing the number of frames $F$ per block increases $V_T$ while reduces $r$. Realize that rates $r_S$ and $r_V$ are dependent variables on $F/T$, while $r_C$ depends only on the camera configuration.
B. OCC-KRF computational complexity

The complexity of OCC-KRF is based mainly on the SVD operations employed in Step 2 and Step 3 (Alg. 1). For the SVD-based pseudo-inversion of $H$ of dimensions $MN \times JL$, its complexity in floating-point operations $O(.)$ is approximately [23]:

$$O_{LS} \triangleq O\{ \max(MN, JL)^2 \cdot \min(MN, JL) \}. \quad (26)$$

As the complexity of the $JL$ SVD calculations of $F \times K$ matrices in Step 3 is

$$O_{SVD} \triangleq O\{ JL \cdot \max(F, K)^2 \cdot \min(F, K) \}, \quad (27)$$

then the OCC-KRF’s complexity is $O_{KRF} = O_{LS} + O_{SVD}$. This complexity may follow a cubic function of the camera resolution, although the data volume $V_T$ in (24) grows only linearly with $MN$.

The screen’s resolution $JL$ appears in both (26) and (27), with a heavier weight in the former. It is evident the process of least-squares restoration in Step 2 of Alg. 1 is more cumbersome than the SVD factorization in Step 3. However, with a time-invariant or very slow degradation in (11), then $(H^T)^\dagger$ can be calculated only once for several transmission blocks.

VI. SIMULATIONS

In this section, numerical analyses demonstrate the proposed system’s functionality and predict its behavior in terms of some of its parameters. Due to the inexistence of tensor-based works for Screen-to-Camera communications, the proposed system is compared to classical matrix-based schemes: in Section VI-B, the video restoration performance is contrasted with the classic degradation matrix inversion process, while in Section VI-C, symbol detection performance is measured against the conventional S2C approach. The simulations are done around the Symbol Error Rate (SER), as well as the video reconstruction error, called here Video Normalized Mean Square Error (Video NMSE) and given by

$$\text{Video NMSE} = \frac{E\{ \|X - \hat{X}\|_F^2 \}}{E\{ \|X\|_F^2 \}}, \quad (28)$$
where $\hat{X}$ is estimated by the OCC-KRF receiver. The operator $E\{\cdot\}$ means the statistical expectation, and $\|\cdot\|_F^2$ indicates the Frobenius norm. Average values are obtained from $10^3$ Monte Carlo runs. For every run, the additive white Gaussian noise power is adjusted to the proper Signal-to-Noise Ratio (SNR) using (14), i.e., $\text{SNR} = \frac{\|Y\|_F^2}{\|N\|_F^2}$. The Frobenius norm of a tensor is defined in [25].

Unless noted otherwise, the degradation matrix (including decimation) is set to $H = I_{MN}$, with $JL = MN$, which certainly agrees with Proposition 1 in Section II. When synthetic videos are used, the entries of $X$ are independently drawn from a continuous Uniform distribution $U(0,A)$, with $A = 1$ representing the highest possible pixel intensity. In general, symbols are generated using the 2-PPM modulation described in Section IV. Section VI-D adresses the PWM and OOK modulations.

### A. Video resolution $J \times L$

Fig. 6 shows the Symbol NMSE and Video NMSE curves for different screen resolutions. The Video NMSE is given by (28) and the Symbol NMSE by

$$\text{Symbol NMSE} = \frac{E\{\|S - \hat{S}\|_F^2\}}{E\{\|S\|_F^2\}}. \quad (29)$$

Indeed, by Fig. 6, screen resolution has no impact on symbol or video estimation error performances. The fact that the increase in resolution also increases the number of unknown elements in $S$ and $X$ accounts for the absent benefit in terms of estimation gain. On the other hand, these parameters impact the data rate and complexity of the OCC-KRF, as shown in Sections V-A and V-B.

### B. Symbol pulses $K$

Fig. 7 compares different SER curves for different $S$ and $P$ values. All values of $K = SP$ agree with condition (35) of Proposition 3 (in the Appendix). As from (25) the symbol rate $r$
may be given by $r = JLSr_V/F$, in Fig. 7 it ranges from $r = 500r_V$ (for $S = 25$) to $r = 2000r_V$ (for $S = 100$). If a video of $r_V = 30$ fps is OCC-encoded as by the method presented in this paper, then the symbol rate in Fig. 7 ranges from 15 kbps to 60 kbps.

From varying $P$, one may note that it brings a coding gain, mainly at a low SNR regime, whereas increasing $S$ does not. This insensitivity of SER to the increase in $S$ is evident from the superposition of the curves for $P = 4$. It is convenient to verify by definition that both $S$ and $P$ determine the size of the matrix $S$ and the third dimension of the tensor $\mathcal{Y}$, but only the former increases the number of symbols to be estimated.

Fig. 8 brings the video NMSE of the simulations done in Fig. 7. In addition, the curves for $(S, P) = (100, 16)$ is displayed. This figure also compares the OCC-KRF curves to the uncoded video restoration scenario. As uncoded one shall understand (17) with $S = 1_{1 \times JL}$, that yields

$$Y^{(3)} = HX^T \in C^{MN \times F},$$

(30)
which is nothing but the conventional non-OCC video degradation model [38]. Then, the video NMSE curve for uncoded video estimation in Fig. 8 is obtained from the following restoration inversion process

$$\hat{X} = \left( H^T \hat{Y}^{(3)} \right)^T,$$

which for simulation in Fig. 8 leads to $$\hat{X} = \left( \hat{Y}^{(3)} \right)^T$$ if $$S = P = 1$$.

Both $$P$$ and $$S$$ impact video restoration performance. As $$K = SP$$ gives the number of rows of $$S$$, an improvement seen with increasing $$K$$ was expected, primarily since the video matrix $$X$$ does not depend on any of these parameters.

From a communication point of view, having a large $$S$$ value, as in Fig. 7 and Fig. 8, is justified only by a higher symbol rate. However, increasing $$S$$ from a certain point does not bring benefits in symbol estimation (Fig. 7) or significant gains in video estimation (Fig. 8), which can be achieved by raising $$P$$ in both cases. In the simulated range of Fig. 8, doubling $$P$$ led to a consistent 3 dB gain in SNR, while doubling $$S$$ from 50 to 100 for $$P = 4$$ only led to a marginal and negligible benefit.
When comparing the performance of the OCC-KRF with the simple inversion of degradation matrix (i.e., $S = P = 1$), it is clear that the proposed tensor-coding scheme effectively enhances the quality of the recovered video. The price to pay is mainly in the increase in the camera’s capture rate and the volume of data generated – see (23) and (24).

C. Frame length $F$

Fig. 9 compares our proposed tensor-based scheme to a matrix-based approach. As conventional OCC matrix-based receivers usually estimate symbols on a frame by frame basis, if in (16) one makes $F = 1$, then one obtains the following matrix formulation

$$\bar{Y} \triangleq HS^T \in \mathbb{C}^{MN \times K}.$$  \hfill (32)

In other words, the third-order tensor $Y \in \mathbb{C}^{MN \times F \times K}$ is reduced to the matrix (32). As for the simulations here $H = I_{MN}$, then simply $\bar{Y} = S^T$. This is consistent with the conventional non-S2C OCC schemes [1], in which the symbols are directly mapped to the screen’s pixel intensities at each frame.
Considering from (25) that symbol rate is given by \( r = J L r_V \alpha \), where \( \alpha = S/F \), we evaluate here two scenarios: a) only \( F \) is increased, and then \( r \) and \( \alpha \) decrease; and b) \( F \) and \( S \) are increased proportionally, then both \( r \) and \( \alpha \) are kept constant. In both scenarios, \( J \), \( L \), and \( r_V \) are constant.

The first thing to note in Fig. 9 is that \( F \) brings a diversity gain, not achieved with \( P \) or \( S \) in Fig. 7.

As the number of unknowns \( S \) does not depend on \( F \), then the increase of this parameter brought the referred estimation gain. This behavior was expected, typical of other communication schemes that use the PARAFAC model. It helps that the entries of \( X \) are random, of a continuous uniform distribution, then this matrix naturally has a conditioning number. As the data rate is inversely proportional to \( F \), a trade-off between diversity gain and data rate exists, similar to countless other space-time block coding schemes.

Now, when the data rate must be kept constant, the results in Fig. 9 indicate that it may be convenient to compensate for the larger frame length by increasing the number of symbols per
block. The same diversity gain with $F$ in Fig. 9 can be found, but the negative effect of the rate reduction was avoided. Negatively, the increase in $S$ impacts both the volume of data generated and the complexity of the receiver. In addition, this increase also requires a proportional increase in the camera’s capture rate $r_C$, as mentioned in Sections V-A and V-B.

In terms of the video NMSE, the impact of $F$ (and $S$) is the opposite of Fig. 9. While diversity gains are attributed only to the increase of $F$ in this figure, this parameter does not improve the video estimation in Fig. 10. Once again, this behavior was expected, as the number of unknowns symbols depend on $S$, while the length of the video is a function of $F$. In both Fig. 9 and Fig. 10, the best estimation scenario occurred when $F$ and $S$ increased proportionally, keeping $\alpha$ constant.

D. Comparison between modulations

The three modulations presented in Section IV serve to demonstrate the flexibility of the proposed transmission scheme for different pulse modulations.
Comparing video restoration performance using the three suggested modulations, Fig. 11 demonstrates that PWM and PPM modulations have similar performance – and are superior to OOK modulation. Given the hypothesis that the generated and simulated symbols within an alphabet are equiprobable, there is a greater presence of zeros in the symbol matrix with the OOK modulation. This sparsity in $S$ is likely the culprit for a worse performance of OOK compared to PWM and PPM.

Figure 11: Video NMSE for OOK, PWM and PPM

E. Real videos

The performance of the proposed system for real videos were also verified by numerical analysis. The results and conclusions found were very similar to those seen in the last simulation results. They are omitted here for the sake of conciseness only.

As an example, Fig. 12 shows the same indexed video frame, in different stages of the transmission process and at different SNR levels. The simulation parameters were $P = 4$, $S = 4$, $F = 80$, $J = M = 72$, and $L = N = 128$. The full video, for different values of SNR, is
Figure 12: (a) uncoded video at the transmitter side; (b) restored at $\text{SNR} = 0$ dB; (c) $\text{SNR} = 10$ dB; (d) $\text{SNR} = 20$ dB; (e) $\text{SNR} = 30$ dB; and (f) $\text{SNR} = 40$ dB.

accessed by https://www.youtube.com/watch?v=OuOVaucpV3Q. Other simulated videos, with different parameters, can be found in the same Youtube channel.

F. Overall discussion

The screen resolution is proportional to the maximum achievable symbol rate at the same time it produces more data to be processed by the receiver. This conclusion is valid for basically all
OCC systems. However, there are no symbol or video estimation improvements when increasing this spatial resolution. It is noteworthy that for now, the screen resolution favors multiplexing gains over spatial diversity gains, which could be obtained if a set of pixels could provide redundancy. The cubic-order complexity of OCC-KRF shall not be neglected for high-resolution screens.

It was observed that increasing symbol pulses $K$ leads to estimation gains in both video reconstruction and symbol estimation. If the number of $S$ per video block is increased, there is a beneficial gain in terms of symbol rate but a marginal SNR gain on video restoration and no gain for symbol detection. On the other hand, the number of pulses per symbol $P$ brings SNR gain for both symbol and video estimation. Besides, a greater $P$ can be used to provide room for dimming control [17], to improve synchronization [42] or to allow even more flexibility to implement other pulse modulations.

Finally, the increase in the number of video frames $F$ per block brought gains on video restoration and symbol detection. The drawback is it also reduces the symbol rate. However, the diversity gains generated by large $F$ may justify the rate reduction, especially if image degradation is critical.

**VII. Conclusion**

This work presented a new Screen-to-Camera (S2C) Optical Camera Communication (OCC) scheme. The novelty is in tensor modeling, using PARAFAC decomposition to model symbol, video, and image degradation as its matrix factors. This approach, never used before for OCC communications, allow the development of a blind receiver capable of restoring the quality of the transmitted video while recovering the encoded symbols. Proposed design rules and numerical analyses allowed us to identify system parameters that can dictate performance in different transmission scenarios. We discussed and weighed the simulation results in light of other performance metrics, such as the symbol rate, the volume of data generated, and the
computational complexity of the proposed algorithm.

Interesting perspectives emerge from this work. For instance, different tensor decompositions in the literature can be adapted to the same framework to provide even better outcomes, more relaxed uniqueness conditions, and lighter estimation algorithms. In addition, image restoration can be further enhanced if the degradation model can be estimated in the proposed model.

APPENDIX

In this section, it is suggested design rules for the proposed algorithms. Precisely, Proposition 1 refers to identifiability conditions for OCC-KRF, whereas the uniqueness of the PARAFAC decomposition (13), in terms of Theorem 1, is approached by Proposition 2 and Proposition 3.

**Proposition 1.** Assume that the degradation matrix $H$ is a full-rank matrix. A sufficient and necessary condition for identifiability of $S$ and $X$ using the OCC-KRF algorithm is that $MN \geq JL$.

**Proof:** The proof of Proposition 1 is straightforward. For OCC-KRF to work as intended, one needs that $(H^T)^\dagger$ be also known, such that degradation can be mitigated, i.e., $H^T (H^T)^\dagger = I_{JL}$ (Step 2 in Alg. 1). Once $H$ is a full-rank matrix by assumption, in order that $H^T (H^T)^\dagger = I_{JL}$, the number of rows of $H$ must be greater than its number of columns, i.e., $MN \geq JL$.

Proposition 1 suggests that using the OCC-KRF algorithm is linked to a proper resolution of the IS at the receiver. In other words, the Region Of Interest (ROI) (in pixels) of the captured video must be at least the size of the screen resolution. Mathematically, this hypothesis is ordinary if the spatial decimation in the image degradation model (11) is disregarded.

Overall, for Proposition 1, the hypothesis of the full rank of $H$ is generically true for Linear Time-Invariant (LTI) degradation models. For LTI systems, $H$ is typically a doubly-block Toeplitz (or Circulant) matrix, built from the elements of an image Kernel (e.g., blurring mask) [43].

**Proposition 2.** Assume that the $H$ is a full-rank matrix, with $MN \geq JL$. Also, assume that
S and X are random matrices with full rank, and that $F \gg 1$ and $JL \gg 1$. Under these hypotheses, the essential uniqueness of the model (13) is almost surely if

$$K \geq \left\lceil \frac{1}{2} + \sqrt{\frac{JL}{F}} \right\rceil. \quad \text{(33)}$$

This is a sufficient, but not necessary, condition.

**Proof:**

If one of the matrix factors of (13) has a full column rank, and other factors are random, then a more relaxed sufficient uniqueness condition than Kruskal’s condition in Theorem 2 can be used. Therefore, from the hypothesis that $H$ has a full-column rank, and that $S$ and $X$ are random, Theorem 3 with correspondence $(A^{(1)}, A^{(2)}, A^{(3)}, I_1, I_2, I_3, R) \iff (H, X, S, MN, F, K, JL)$ leads to the following sufficient uniqueness condition

$$K(K - 1)F(F - 1) \geq 2JL(JL - 1). \quad \text{(34)}$$

In [34] it is emphasized that (34) might be used even if $S$ or $H$ were not random. Finally, considering the statement that $F \gg 1$ and $JP \gg 1$, then (34) becomes $K^2 - K \geq 2 \left( \frac{JL}{F} \right)^2$, whose only integer positive solution is given by (33), which ends this proof.  

**Proposition 3.** Assume the hypotheses of Proposition 2 and that $\min(K, F) \geq 3$, then the decomposition (13) is essentially unique, almost surely, if

$$K \geq \left\lceil \frac{JL}{F} + 1 \right\rceil. \quad \text{(35)}$$

**Proof:** Consider the same correspondence $(A^{(1)}, A^{(2)}, A^{(3)}, I_1, I_2, I_3, R) \iff (H, X, S, MN, F, K, JL)$. Also, let $MN \geq JL$ and $\min(F, K) \geq 3$, as stated. From Theorem 4 of Section II, a sufficient condition to guarantee the uniqueness of the PARAFAC model is

$$(F - 1)(K - 1) \geq MN. \quad \text{(36)}$$
As $F \gg 1$ by hypothesis stated in Proposition 2, then (36) becomes 35, ending the proof.

The conditions on $K$ present in Proposition 2 are more relaxed than those presented in Proposition 3. In both cases, screens with higher resolutions force an increase in the number of frames captured by the optical camera.

REFERENCES

[1] “IEEE standard for local and metropolitan area networks—part 15.7: Short-range optical wireless communications,” IEEE Std 802.15.7-2018 (Revision of IEEE Std 802.15.7-2011), pp. 1–407, 2019.
[2] R. Boubezari, H. Le Minh, Z. Ghassemlooy, and A. Bouridane, “Smartphone camera based visible light communication,” J. Lightw. Technol., vol. 34, no. 17, pp. 4121–4127, 2016.
[3] M. D. Thieu, T. L. Pham, T. Nguyen, and Y. M. Jang, “Optical-RoI-signaling for vehicular communications,” IEEE Access, vol. 7, pp. 69873–69891, 2019.
[4] T. Nguyen, A. Islam, and Y. M. Jang, “Region-of-interest signaling vehicular system using optical camera communications,” IEEE Photon. J., vol. 9, no. 1, pp. 1–20, 2017.
[5] Y. Li, Z. Ghassemlooy, X. Tang, B. Lin, and Y. Zhang, “A VLC smartphone camera based indoor positioning system,” IEEE Photon. Technol. Lett., vol. 30, no. 13, pp. 1171–1174, 2018.
[6] D. T. Nguyen, S. Park, Y. Chae, and Y. Park, “VLC/OCC hybrid optical wireless systems for versatile indoor applications,” IEEE Access, vol. 7, pp. 22371–22376, 2019.
[7] R. D. Roberts, “A MIMO protocol for camera communications (camcom) using undersampled frequency shift ON-OFF keying (UFSOOK),” in 2013 IEEE Globecom Workshops (GC Wkshps), 2013, pp. 1052–1057.
[8] N. Saeed, S. Guo, K.-H. Park, T. Y. Al-Naffouri, and M.-S. Alouini, “Optical camera communications: Survey, use cases, challenges, and future trends,” Physical Communication, vol. 37, p. 100900, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1874490719304859
[9] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, “6G wireless networks: Vision, requirements, architecture, and key technologies,” IEEE Veh. Technol. Mag., vol. 14, no. 3, pp. 28–41, 2019.
[10] M. Z. Chowdhury, M. Shahjalal, S. Ahmed, and Y. M. Jang, “6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions,” IEEE Open J. Commun. Soc., vol. 1, pp. 957–975, 2020.
[11] J. Klein, J. Xu, C. Brauers, J. Jochims, and R. Kays, “Investigations on temporal sampling and patternless frame recovery for asynchronous display-camera communication,” IEEE Trans. Circuits Syst. Video Technol., pp. 1–1, 2021.
[12] R. Kays, C. Brauers, and J. Klein, “Modulation concepts for high-rate display-camera data transmission,” in 2017 IEEE Int. Conf. on Communications (ICC), 2017, pp. 1–6.
[13] T. Fujihashi, T. Koike-Akino, P. V. Orlik, and T. Watanabe, “High-throughput visual MIMO systems for screen-camera communications,” IEEE Trans. Mob. Comput., vol. 20, no. 6, pp. 2200–2211, 2021.
[14] Cisco, “Cisco annual internet report (2018–2023) white paper,” Tech. Rep., March 2020.

[15] Y. Yang, J. Hao, and J. Luo, “Ceilingtalk: Lightweight indoor broadcast through LED-camera communication,” *IEEE Trans. on Mobile Computing*, vol. 16, no. 12, pp. 3308–3319, 2017.

[16] D. Vuong and M. Yoo, “Interpixel interference mitigation in visible light communication using image sensor,” *IEEE Access*, vol. 6, pp. 45 543–45 551, 2018.

[17] Z. Wang, Q. Wang, W. Huang, and Z. Xu, *Optical Camera Communication: Modulation and System Design*, 2018, pp. 291–351.

[18] T. Nguyen, A. Islam, T. Yamazato, and Y. M. Jang, “Technical issues on IEEE 802.15.7m image sensor communication standardization,” *IEEE Communications Mag.*, vol. 56, no. 2, pp. 213–218, 2018.

[19] 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 Photon. J.*, vol. 7, no. 5, pp. 1–12, 2015.

[20] P. Luo, M. Zhang, Z. Ghassemlooy, H. Le Minh, H.-M. Tsai, X. Tang, and D. Han, “Experimental demonstration of a 1024-QAM optical camera communication system,” *IEEE Photon. Technol. Lett.*, vol. 28, no. 2, pp. 139–142, 2016.

[21] R. D. Roberts, “Undersampled frequency shift ON-OFF keying (UFSOOK) for camera communications (camcom),” in *2013 22nd Wireless and Optical Communication Conf.*., 2013, pp. 645–648.

[22] A. Wang, Z. Li, C. Peng, G. Shen, G. Fang, and B. Zeng, “Inframe++: Achieve simultaneous screen-human viewing and hidden screen-camera communication,” in *Proc. of the 13th Ann. Int. Conf. on Mobile Systems, Applicat., and Services*, ser. MobiSys ’15. New York, NY, USA: Assoc. for Computing Machinery, 2015, p. 181–195. [Online]. Available: https://doi.org/10.1145/2742647.2742652

[23] L. R. Ximenes, G. Favier, and A. L. F. de Almeida, “Closed-form semi-blind receiver for MIMO relay systems using double Khatri-Rao space-time coding,” *IEEE Signal Process. Lett.*, vol. 23, no. 3, pp. 316–320, 2016.

[24] W. d. C. Freitas, G. Favier, and A. L. F. de Almeida, “Sequential closed-form semiblind receiver for space-time coded multihop relaying systems,” *IEEE Signal Process. Lett.*, vol. 24, no. 12, pp. 1773–1777, 2017.

[25] N. D. Sidiropoulos, L. De Lathauer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, “Tensor decomposition for signal processing and machine learning,” *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, 2017.

[26] A. Cichocki, “Tensor networks for big data analytics and large-scale optimization problems,” *CoRR*, vol. abs/1407.3124, 2014. [Online]. Available: http://arxiv.org/abs/1407.3124

[27] L. Kuang, F. Hao, L. T. Yang, M. Lin, C. Luo, and G. Min, “A tensor-based approach for big data representation and dimensionality reduction,” *IEEE Trans. Emerg. Topics Comput.*, vol. 2, no. 3, pp. 280–291, 2014.

[28] L. Kassab, A. Kryshchenko, H. Lyu, D. Molitor, D. Needell, and E. Rebrova, “On nonnegative matrix and tensor decompositions for COVID-19 twitter dynamics,” *CoRR*, vol. abs/2010.01600, 2020. [Online]. Available: https://arxiv.org/abs/2010.01600

[29] Y.-h. Taguchi, S. A. P. Dharshini, and M. M. Gromiha, “Identification of transcription factors, biological pathways, and
diseases as mediated by N6-methyladenosine using tensor decomposition-based unsupervised feature extraction,” *Appl. Sciences*, vol. 11, no. 1, 2021. [Online]. Available: https://www.mdpi.com/2076-3417/11/1/213

[30] R. Harshman, “Foundations of the PARAFAC procedure: Models and conditions for an “explanatory” multi-modal factor analysis,” *UCLA Working Papers in Phonetics*, vol. 16, 1970.

[31] N. D. Sidiropoulos and R. S. Budamati, “Khatri-Rao space-time codes,” *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2396–2407, 2002.

[32] J. B. Kruskal, “Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics,” *Linear Algebra and its Applicat.*, vol. 18, no. 2, pp. 95–138, 1977. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0024379577900696

[33] L. De Lathauwer, “A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization,” *SIAM J. on Matrix Anal. and Applicat.*, vol. 28, no. 3, pp. 642–666, 2006. [Online]. Available: https://doi.org/10.1137/040608830

[34] A. Stegeman, “On uniqueness conditions for Candecomp/PARAFAC and Indscal with full column rank in one mode,” *Linear Algebra and its Applicat.*, vol. 431, no. 1, pp. 211–227, 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0024379509001104

[35] L. Chiantini and G. Ottaviani, “On generic identifiability of 3-tensors of small rank,” *SIAM Journal on Matrix Analysis and Applicat.*, vol. 33, no. 3, pp. 1018–1037, 2012.

[36] I. Domanov and L. D. Lathauwer, “Generic uniqueness conditions for the canonical polyadic decomposition and indscal,” *SIAM Journal on Matrix Analysis and Applicat.*, vol. 36, no. 4, p. 1567–1589, Jan 2015. [Online]. Available: http://dx.doi.org/10.1137/140970276

[37] V. Strassen, “Rank and optimal computation of generic tensors,” *Linear Algebra and its Applicat.*, vol. 52-53, pp. 645–685, 1983. [Online]. Available: https://www.sciencedirect.com/science/article/pii/002437958380041X

[38] R. C. Gonzalez and R. E. Woods, *Digital image processing*. USA: Prentice-Hall, Inc., 2008.

[39] P. Luo, Z. Ghassemlooy, H. Le Minh, X. Tang, and H.-M. Tsai, “Undersampled phase shift ON-OFF keying for camera communication,” in *2014 Sixth Int. Conf. on Wireless Communications and Signal Processing (WCSP)*, 2014, pp. 1–6.

[40] S. H. Lee, S.-Y. Jung, and J. K. Kwon, “Modulation and coding for dimmable visible light communication,” *IEEE Communications Mag.*, vol. 53, no. 2, pp. 136–143, 2015.

[41] A. Wilkins, J. Veitch, and B. Lehman, “LED lighting flicker and potential health concerns: IEEE standard PAR1789 update,” in *2010 IEEE Energy Conversion Congress and Exposition*, 2010, pp. 171–178.

[42] Y. Shiraki, T. G. Sato, Y. Kamamoto, and T. Moriya, “Flexible synchronization in optical camera communication with On-Off keying,” in *2017 IEEE Globecom Workshops (GC Wkshps)*, 2017, pp. 1–6.

[43] A. K. Jain, *Fundamentals of Digital Image Processing*. USA: Prentice-Hall, Inc., 1989.