Composite Amplitude-Shift Keying for Effective LED-Camera VLC

Yanbing Yang  Jun Luo

Abstract—LED-Camera Visible Light Communication (VLC) is gaining increasing attention, thanks to its readiness to be implemented with Commercial Off-The-Shelf devices and its potential to deliver pervasive data services indoors. Nevertheless, existing LED-Camera VLC systems employ mainly low-order modulations such as On-Off Keying (OOK) given the simplicity of their implementation, yet such rudimentary modulations cannot yield a high throughput. In this paper, we investigate various opportunities of using a high-order modulation to boost the throughput of LED-Camera VLC systems, and we decide that Amplitude-Shift Keying (ASK) is the most suitable scheme given the limited operating frequency of such systems. However, directly driving an LED to emit different levels of luminance may suffer heavy distortions caused by the nonlinear behavior of LED. As a result, we innovatively propose to generate ASK using the composition of light emission. In other words, we digitally control the On-Off states of several groups of LED chips, so that their light emissions compose in the air to produce various ASK symbols. We build a prototype of this novel ASK-based VLC system and demonstrate its superior performance over existing systems: it achieves a rate of 2kbps at a 1 m distance with only a single LED luminaire for static users and more than 1kbps for mobile users.

Index Terms—Visible Light Communication, Mobile Computing, Collaborative Transmissions; Grayscale Modulation

1 INTRODUCTION

While WiFi is becoming more and more pervasive, using it as an information broadcast service (e.g., delivering notices, advertisements, or even emergency alerts) is still not an option mainly due to its relatively high overhead and lack of location awareness. As an alternative, high-speed Visible Light Communication (VLC) based on Light Emitting Diodes (LEDs) and Photo-Diodes (PDs) [1], [2], [3], [4], [5] is yet to see its avatar after decades of theoretical studies, since it is severely challenged by the interference in a real-life scenario. Consequently, LED-Camera VLC [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], with its sole reliance on Commercial-Off-The-Shelf (COTS) devices, rises as a readily deployable service, strongly backed by the ever-growing popularity of camera-equipped smartphones. Moreover, using camera as a receiver can largely eliminate interference suffered by LED-PD VLC [23], thanks to the inherent spatial division at the receiver [11]. Nevertheless, the limited frequency response of camera makes it hard to gain a sufficiently high data rate.

Various seminal approaches have been taken to improve LED-Camera VLC with respect to both throughput and reliability in the past few years [8], [9], [10], [11]. Earlier proposals such as [6], [24] make use of reflected light as the communication media and adopt On-Off Keying (OOK) as the basic modulation scheme, so the achievable data rate is reported to be only a few bytes per second. Recent proposals manage to raise the data rate to kbps level by switching to direct light as communication media [9], engineering advanced modulation schemes such as Color-Shift Keying (CSK) [10], [25], and applying sophisticated channel coding such as rateless codes [11]. However, these systems all share a bottleneck on the receiver side, i.e., the smartphone camera, although an LED transmitter can be modulated at a frequency up to GHz. Basically, the rolling shutter effect of a CMOS camera exploited by LED-Camera VLC offers a frequency response only up to a couple tens of kHz [9], [11], leading to a relatively low optical clock rate confining the performance of such a system. Moreover, the nature of rolling shutter also limits usable modulation schemes, preventing more advanced modulation techniques such as phase-shift from being applied. According to the IEEE standard for VLC [26], the data rate of a practical VLC system should be at least tens of kbps, so there is still a big gap between existing LED-Camera VLC systems and an applicable VLC data service.

One of the major reasons causing a low data rate of existing LED-Camera VLC is the commonly applied low-order modulation, such as OOK. Several existing works attempted to leverage high-order modulations for boosting data rates. For example, ColorBars [10] and undersampled CSK [25] adopt CSK inspired by Wavelength-Division Multiplexing (WDM), yet the resulting low signal strength has reduced the transmission distance to a few centimeters. ReflexCode [27] proposes Grayscale-Shift Keying (GSK) by exploring the collaborative transmissions from three adjacent light sources, but it may face problems if only one luminaire is available. UPAM [13] devises an undersampled modulation scheme based on multiple amplitude modulation (MAM) to maintain non-flicking transmission in LED-Camera VLC, but it only achieves a low data rate of 250 bps.
In fact, there is even no systematic study on what modulation schemes can be applied for LED-Camera VLC by far; OOK is commonly applied only due to its simplicity. In this paper, we make an attempt to improve the data rate of LED-Camera VLC by first answering the following question: what are the feasible modulation schemes under a low optical clock rate? Whereas the answer reveals (as expected) Amplitude-Shift Keying (ASK) as the only choice, it also suggests that a straightforward implementation of ASK by digitally driving the luminance of LEDs can suffer severe distortion induced by the nonlinear behavior of an LED. Moreover, demodulating a high-order ASK under kHz clock rate is non-trivial because of the reduced symbol distance compared with OOK. To this end, we set out to devise a new modulation mechanism termed Composite Amplitude-Shift Keying (CASK) for LED-Camera VLC. CASK controls the ON-OFF states of several groups of LED chips individually, so that the composite light emission exhibits various levels of grayscale (i.e., amplitude-shift). Also, CASK employs a delicate demodulation algorithm to properly recognize grayscale symbols under a kHz-level clock rate. In summary, our main contributions are as follows:

- We systematically investigate the applicability of various modulation schemes under LED-Camera VLC.
- We propose the novel idea of generating ASK through physical light composition without using sophisticated transmitter circuits, in order to overcome LED nonlinear distortion.
- We engineer an efficient demodulation algorithm to cope with the reduced symbol distance under a high-order ASK.
- We set up a communication model for analyzing the communication capacity of LED-Camera VLC using high order modulation schemes.
- We build CASK into a practical VLC prototype; it adopts COTS LEDs as the transmitter and an Android phone as the receiver.
- We conduct extensive evaluations on this prototype to demonstrate the efficacy of CASK in boosting the throughput of LED-Camera VLC for both static and mobile users.

We focus only on using a single COTS LED luminaire as the transmitter in this paper. Applying rateless codes to combine several transmitters in a collaborative transmission manner has been proposed in [11], and multiple transmitters operating CASK should be compatible with such a collaborative extension for further improving VLC throughput. Moreover, optimized LED layout techniques [5] and advanced handover methods [31], [32] could be utilized to further enhance the performance of CASK-enabled VLC systems under user mobility.

## 2 Preliminary Study and Motivation

This section serves as a research on the feasibility of various modulation schemes for LED-Camera VLC (thus also a motivation for our work), as well as a literature survey on recent proposals for such systems. We also summarize the performance comparisons between the proposed CASK and the existing proposals in Table 1.

### 2.1 Rolling Shutter and OOK Basics

Rolling shutter is a special property of CMOS cameras, i.e., the pixels of a single frame are not exposed and sampled together but rather sequentially in a column-by-column manner. As a result, we may deem the columns exposed at the same time a sampler to a time-varying process in terms of certain lighting property (e.g., luminance and tone) [6]. As a special case, OOK modulates input signal onto the ON-OFF states of an LED, and the resulting bright-dark luminance process is sampled by a CMOS camera (via its rolling shutter) as bright-dark bands in a frame. We omit the detailed illustration due to page limit, but rather refer the readers to Figure 3 of [9]. Here we only use a frame sampled by our CASK receiver, see Fig. 1(a), to assist in highlighting a few key points.

First of all, rolling shutter has two crucial parameters, namely rolling-shutter frequency $F_r$ and scanning (column) width $W_r$ [11]. If a modulation process generates symbols at frequency $F_m$, the necessary condition for an LED-Camera VLC system to work is $F_m \leq F_r$. Although $F_r$ may vary with different smartphones, it is normally at the level of 20 kHz. Therefore, given that $W_r$ is only a couple of pixels and the communication is asynchronous (i.e., the rolling shutter sampling may not be aligned with the modulation process), $F_m \leq 10 \text{kHz}$ is needed to guarantee a measurable symbol width in a frame. Secondly, though it is feasible to

<table>
<thead>
<tr>
<th>Summary of Camera-based VLC Systems.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Capacity</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Seminal [6]</td>
</tr>
<tr>
<td>HybridVLIC [7]</td>
</tr>
<tr>
<td>RollingLight [28]</td>
</tr>
<tr>
<td>ColorBars [10]</td>
</tr>
<tr>
<td>[13]</td>
</tr>
<tr>
<td>[25]</td>
</tr>
<tr>
<td>CeilingTalk [29]</td>
</tr>
<tr>
<td>POLI [30]</td>
</tr>
<tr>
<td>CASK</td>
</tr>
</tbody>
</table>

![Fig. 1. Illustrating rolling shutter and modulations. (a) When setting $F_m = F_r$, the time lapse between sampling two brands is exactly $1/F_r$, and the width of each band is $W_r$. (b) BFSK constructed by two OOKs with different frequencies. (c) A sine wave sampled by rolling shutter.](image-url)
have $F_m \ll F_r$, the resulting symbol width grows as $F_m$ decreases. Since the number of symbols contained in a frame is bounded by the minimum between the screen width and the width of the transmitter’s image (a.k.a., Region of Interest, or RoI [33]), widening individual symbols can only reduce the symbol rate (hence the throughput), given a fixed frame rate of 30fps [11].

Last but not least, as the transmitter piggybacks on an existing light infrastructure, it is necessary that the modulation process does not generate visible flickers on the individual LED luminaires. This demands that the modulation process should not produce low frequency components. In other words, $F_m$ is bounded below by a couple of kHz. While the limited operating frequency of rolling shutter receivers confines the selection of modulation schemes (as will be discussed soon), their one-dimension sampling nature rules out the feasibility of the two-dimensional bar codes like coding mechanism commonly used by Screen-Camera VLC [16], [17], [18], [20].

2.2 FSK and PWM Are Not Beneficial

Frequency-Shift Keying (FSK), as a very conventional RF modulation scheme, was adopted by RollingLight [9], a seminal proposal that introduced direct light to LED-Camera VLC. Basically, an FSK symbol consists of several bright-dark bands caused by an OOK modulation running at a given frequency, shown in Fig. 1(b). Apparently, this scheme could substantially reduce the symbol rate as basic OOK only requires one band per symbol. Moreover, FSK has to run at frequencies lower than $F_m$ (which is already below 10 kHz) in order to produce different symbols. As a result, FSK may only produce a data rate of several bytes per second [9].

If we push the limit of FSK to the extent of only two bands per symbol, we end up with Pulse-Width Modulation (PWM). However, ideal PWM is not feasible given the limit of rolling shutter. As we mentioned early, rolling shutter samples a time-varying process with a granularity of rolling shutter. As we mentioned early, rolling shutter samples a time-varying process with a granularity of rolling shutter certainly cannot match.

Fig. 2. LED nonlinearity distorts 8-ASK symbols.

2.3 Waveform Should Be Avoided in Modulation

Phase-Shift Keying (PSK) related modulation schemes (in particular Quadrature Amplitude Modulation, or QAM) are widely used in RF communications [34]. However, all such methods rely on sampling complicated waveforms at the receiver with pulse sequences of some very high frequency (notably at GHz level). According to our earlier analysis on rolling shutter, its sampling ability is far from sufficient to handle waveforms, unless we allow a super wide waveform/symbol to be produced, as shown by Fig. 1(c), thus ruining the data rate. For the same reason, Orthogonal Frequency-Division Multiplexing (OFDM), another popular RF modulation scheme that is envisioned also for LED-PD VLC, is Pulse-Position Modulation (PPM). It is used in Ultra-Wide band (UWB) [37] and is adopted by DarkLight [38] for enabling an LED-PD VLC to operate under light-off scenarios. As the bandwidth required by PPM is extremely high (MHz level for DarkLight), the limited operating frequency of rolling shutter certainly cannot match.

2.4 CSK May Not Scale Up

Inspired by Wavelength-Division Multiplexing (WDM), ColorBars [10] exploits Tri-LED’s ability of producing a wide range of colors to generate a high order modulation termed Color-Shift Keying (CSK). In fact, CSK can also be deemed as a kind of FSK, where the symbol frequency is determined by the light wavelength rather than the duty cycle rate of OOK (as used by RollingLight [9]). As increasing ColorBars’ transmission distance of a few centimeters to meter level requires a high-power LED with condenser cup [25], [12], the key question here is whether CSK scales up to a full lighting infrastructure. A first but relatively minor issue is that Tri-LEDs are more expensive than phosphor-coated white LEDs commonly used for commercial lighting, so it is quite unlikely that we will see large scale adoption of Tri-LEDs in commercial lighting. Secondly but more importantly, applying CSK to COTS LED luminaires requires a close synchronization among tens to hundreds LED chips. This, on one hand, makes the driver extremely complicated, and on the other hand, it results in a rather unreliable modulation whose color symbols are prone to distortion caused by asynchronous light emissions. Thirdly, a slight loss of synchronization among LED chips can also break the balanced emission designed by CSK to avoid flicker, causing visible flicker on the individual LED luminaires.

2.5 ASK Is The Last Choice, But ...

With our aforementioned analysis, ASK appears to be the only remaining choice. An ideal ASK extends OOK by producing grayscale bands between the brightest and the darkest levels, driven digitally by some input signals. Nonetheless, LEDs is well known for their nonlinear behavior [39]; in fact, the same reason has long been one of the major challenges for the development of OFDM-based LED-PD VLC. Essentially, while the given input signal indicates 2 in an 8-ASK, the output at an LED can severely deviate
from 2/7 grayscale (as shown by Fig. 2), and this distortion may vary with different types of LEDs and ambient conditions (e.g., temperature) as well. As a result, existing modulation schemes designed for LED-PD VLC all confine their (LED) output dynamic range to a pseudo-linear section of LED transfer characteristics [36], [40]. Such a makeshift may work for high-power LED-PD VLC, but it is certainly not feasible for supporting ASK given COTS LEDs, as it would significantly reduce the symbol distance, causing a much higher error rate in demodulation. Consequently, novel physical layer techniques have to be in place to handle LED nonlinearity.

3 CASK: MODULATION AND DEMODULATION

In this section, we first elaborate how our novel CASK works on the transmitter side to generate ASK symbols via physical light composition, replacing conventional driver circuits that are both distortion-prone and power-intensive. Then we present our carefully designed demodulation process for recognizing CASK symbols on the receiver side. We omit the descriptions of common blocks in LED-Camera VLC, such as Forward Error Correction (FEC) and RoI extraction [8], [11].

3.1 From Data to Light

The transmitter in a conventional ASK-based VLC system, as shown in Fig. 3(a), first divides a data stream into packets with FEC encoding to combat packet loss. Then the encoded packets are modulated to various ASK symbols, followed by a Digital-to-Analog Converter (DAC) to convert digital signals to analog domain. A Power Amplifier (PA) is finally added to deliver sufficient power for driving the LED transmitter. However, such a circuit may cause significant symbol distortions thanks to both the LED nonlinearity explained in Section 2.5 and the involvement of analog waveform sensitive to noise corruption. In addition, the DAC and PA are often considered as power-intensive components, potentially adding more energy consumption on the lighting infrastructure.

Our idea of CASK generates ASK symbols with only digital control signals as successfully used by OOK and FSK modulations. Inspired by the spatial filtering technique for beamforming in RF technology [41], it is possible to constructively combine the light emissions from several groups of LED chips to form different luminance levels (hence ASK symbols) [42]. Meanwhile, commercial LED luminaires often consist of multiple LED chips. Intuitively, if we separate these chips into several groups and control their ON-OFF process according to the modulated signal, their composite light emission will generate corresponding ASK symbols. Based on this idea, we propose CASK modulation circuit that totally gets rid of the analog part, as shown in Fig. 3(b). We present the details on how CASK produces ASK symbols in the next section.

3.2 Exploiting Composite Light Emission for ASK Symbol Generation

Under conventional OOK, a whole LED luminaire is toggled between ON and OFF states. Such a modulation ignores the nature of commercial LED luminaires that often contain multiple LED chips. Given an LED luminaire consisting of $N$ LED chips and according to how many LED chips are ON at a given moment, the emitted light intensity should have $N + 1$ levels, potentially resulting in $(N + 1)$-ASK with each luminance level representing a symbol. Using such a modulation may boost data rates of an OOK VLC system by $(N + 1)/2$-fold in theory. Apparently, the higher the value of $N$, the more bits can be represented by each symbol,
TABLE 2
ASK symbols produced by different schedules of LED strips.

<table>
<thead>
<tr>
<th>Symbols</th>
<th>ON-OFF schedules of LED strips</th>
</tr>
</thead>
<tbody>
<tr>
<td>OOK</td>
<td>S0</td>
</tr>
<tr>
<td>A0</td>
<td>OFF</td>
</tr>
<tr>
<td>A1</td>
<td>ON</td>
</tr>
<tr>
<td>A2</td>
<td>ON</td>
</tr>
<tr>
<td>A3</td>
<td>ON</td>
</tr>
<tr>
<td>A4</td>
<td>OFF</td>
</tr>
<tr>
<td>A5</td>
<td>OFF</td>
</tr>
<tr>
<td>A6</td>
<td>OFF</td>
</tr>
<tr>
<td>A7</td>
<td>OFF</td>
</tr>
<tr>
<td>4ASK</td>
<td>S0</td>
</tr>
<tr>
<td>A0</td>
<td>OFF</td>
</tr>
<tr>
<td>A1</td>
<td>OFF</td>
</tr>
<tr>
<td>A2</td>
<td>OFF</td>
</tr>
<tr>
<td>A3</td>
<td>OFF</td>
</tr>
<tr>
<td>A4</td>
<td>OFF</td>
</tr>
<tr>
<td>A5</td>
<td>OFF</td>
</tr>
<tr>
<td>A6</td>
<td>OFF</td>
</tr>
<tr>
<td>A7</td>
<td>OFF</td>
</tr>
<tr>
<td>8ASK</td>
<td>S0</td>
</tr>
<tr>
<td>A0</td>
<td>OFF</td>
</tr>
<tr>
<td>A1</td>
<td>OFF</td>
</tr>
<tr>
<td>A2</td>
<td>OFF</td>
</tr>
<tr>
<td>A3</td>
<td>OFF</td>
</tr>
<tr>
<td>A4</td>
<td>OFF</td>
</tr>
<tr>
<td>A5</td>
<td>OFF</td>
</tr>
<tr>
<td>A6</td>
<td>OFF</td>
</tr>
<tr>
<td>A7</td>
<td>OFF</td>
</tr>
</tbody>
</table>

raising the order of the ASK. However, a higher \( N \) also causes a reduced symbol distance, potentially increasing the error rate in demodulation. By far, we have tested up to 8-ASK with a reasonable transmission distance at meter level under our current hardware limit, yet we are on the way to upgrade our prototype in order to support higher modulation orders. Fig. 4 shows the ASK symbols of 2-, 4-, and 8-ASKs generated by our CASK prototype. In reality, an LED luminaire can have more chips than the order of ASK, so we group the chips into strips, and consider each strip as a controllable unit.

In our current prototype as described in 4.1, we employ 7 controllable LED strips forming an LED luminaire (a VLC transmitter), so that it can yield maximum 8-ASK in theory. Here we take 4-ASK as an example to explain how to form specific ASK symbols via composite light emission. As 3 LED strips are enough to produce 4-ASK, we simply group LED strips to reduce the order of ASK. As summarized in Table 2, we switch all LED strips OFF to generate Amplitude-0 (or A0), then the middle two strips (S1 and S2)\(^2\) are turned ON and others OFF to form Amplitude-1 (or A1). Subsequently, the rest 5 LED strips ON and others OFF to generate Amplitude-2 (or A2) (such a symmetrical setting is needed for flicker suppression as analyzed in Section 3.3). Finally, all LED strips are ON for Amplitude-3 (or A3). As generalizing these to higher order ASKs is straightforward, we omit the detailed explanations for 8-ASK but refer readers to Table 2.

3.3 Flicker Suppression with Balance Coding

As LED transmitters piggyback on a lighting infrastructure, transmissions should cause no visible flicker. Unfortunately, directly using CASK would not be flicker-free. For example, a symbol sequence of A0, \cdots, A3, \cdots, A0, \cdots would generate a low frequency component with high power, causing visible flicker. For existing OOK or FSK VLC systems, Run-Length Limited (RLL) codes (as recommended in IEEE 802.15.7 [43]) are commonly used to maintain DC balance and bit disparity, but these codes are not devised for a high-order ASK. Therefore, we propose a new coding scheme inspired by Manchester coding, where each symbol is extended to include itself and its complementary symbol, i.e., a symbol \( A_i \) becomes \( A_iA(3-i) \) for 4-ASK. This coding, on one hand, maintains DC balance, and on the other hand, boils down to Manchester coding for individual LED strips. Therefore, CASK enhanced with our new coding is guaranteed to be flicker-free, and we have conducted user study to confirm it.

3.4 ASK Demodulation on Smartphone Receiver

In an LED-Camera VLC system, the receiver captures sent messages via rolling shutter camera, and the transmitter is projected as banded sections (RoIs) carrying information in a frame. Therefore, the demodulator first extracts all RoIs in a frame\(^3\), and then converts those bands into a grayscale sequence. For conventional OOK modulated signals, a pre-configured threshold can be used to distinguish only two amplitude states: bright (A1) or dark (A0). However, such a trivial threshold configuration is not suitable for CASK, since a higher order ASK reduces symbol distance. This, on one hand, complicates demodulation, and on the other hand, makes a symbol more sensitive to noise and tonal range variation.

Inspired by the preamble setting in CDMA, a few successive brightest symbols is inserted in a packet header. As a result, the header contains the grayscale information

---

2. Other settings either reduce the symbol distance (e.g., only S1 ON) or break the symbol balance (e.g., 3 strips ON for A1 and 5 strips ON for A2).

3. We omit the detailed descriptions for RoI extraction, but rather refer readers to Section 3.3 in CeilingTalk [29] for brevity.
of the whole packet as shown in Fig. 5, and the tonal
range variation across the frame can be further derived by
measuring multiple headers. Consequently, we can measure
the grayscale distribution in a frame so as to set appro-
priate thresholds for demodulating symbols. Specifically,
the demodulator first identifies headers by using a rough
threshold based on tonal range distribution in the previous
demodulation stage. Once all headers in a frame are located,
a packet along with its grayscale envelope is determined
by two consecutive headers as shown in Fig. 5 (EN in the
figure). Based on this envelope, the detection thresholds
for all symbols (i.e., lower bounds in grayscale that are
relative to the envelope) are set empirically to maximize the
chance for correct demodulation; they are inherently stable
because they are produced by physical composition. In the
following, we present the detailed demodulation process
for retrieving data bits from a given packet (a section of a
banded image).

### 3.4.1 Demodulation Based on Width

In a conventional OOK-based system, a symbol is recog-
nized by detecting the width of a band using a bisection
threshold to its grayscale. To combat the blooming effect that
causes bright bands wider than the darker ones, [11] pro-
poses a method by reasoning on the widths of bright/dark
bands in a clustered manner. Nevertheless, a higher order
ASK modulation requires a substantially different demod-
ulation scheme thanks to its reduced symbol distance in
grayscale as shown by Fig. 4. In addition, a higher transmis-
sion frequency further complicates demodulation, because
symbols (bands) around the midtone get their width sig-
nificantly shrunk, as demonstrated by comparing the two
packets transmitted at two different frequencies in Fig. 5.
Therefore, the width-based demodulation alone is not suit-
able anymore as the substantially narrowed bands are much
more prone to the corruption by the blooming effect.

Fig. 5(b) also reveals another issue: a symbol can hardly
get stable before transiting to next symbol at a higher
frequency. This transiency is harmful as grayscale values
cannot be solely relied upon for demodulation. For example,
if two symbols A0 and A7 are sent in a row, there must
exist pixels with a grayscale value corresponding to other
symbols (say A4) between them, and there is no way to
differentiate these “fake symbols” from those becoming
transient due to a high frequency. Consequently, we need
to upgrade the width-based detection and value-based
demodulation certain new scheme that takes both variance
and transiency into consideration.

### 3.4.2 Leveraging the Gradient of Grayscale

Apparently, a feasible approach should first determine
which bands are representing valid symbols, before con-
verting them into data bits. A useful observation drawn
from Fig. 5(b) is that the gradient in grayscale is much
more indicative than the grayscale value itself. Considered
again the aforementioned example, the gradient between
two consecutive symbols A0 and A7 should always be
steeper than having another symbol, say A4, between them,
not matter how transient the band representing that symbol
is. Therefore, our CASK demodulation extracts grayscale
gradient as an additional criterion to be combined with
width and value.

Essentially, we look into the grayscale function \( G(i) \), \( i = 1, 2...n \) of pixel brightness in a frame, where \( n \) is the total
number of horizontal pixels and \( G(\cdot) \) returns the average
pixel brightness for a given column. In an interval between
your two neighboring thresholds, a valid symbol causes a
certain gradient variance (determined by the transmission
frequency) that distinguishes itself from other valid symbols
or no symbol. In particular, \( G'(\cdot) \), the first-order derivative
of \( G(\cdot) \), should approach to zero where a certain symbol
exists. Since \( G(i) \) is discrete grayscale values of given pixels,
the value of \( G'(i) \) may not exactly equal to zero. As a
result, we examine where the value of the finite difference
of \( G(i) \) goes across zero. Basically, **Algorithm 1** first makes
use of \( G'(i) \) crossing zero to identify local minimal and/or
maximal, which represents candidate symbols such as A1
and/or A5 as shown in Fig. 5(b) (marked by a red \( \times \)). In
order to detect potential symbols appearing transiently in-
between two other symbols (e.g., A4 and A3 between A6
and A1, as shown in Fig. 5(b) between the two red dash-
dotted lines), the algorithm further employs the second-
order derivative \( G''(i) \) crossing zero to detect positive
and/or negative inflection points as candidates for such
transient symbols. All these candidate values are put into
\( S_A \), and the demodulation procedure compares them with
the thresholds to convert them into bits.

To summarize, the CASK demodulator first locates all
headers by measuring brightness and width information,
and extracts a potential packet between two consecutive
headers. For each potential packet, all pixels (averaged
per column) will be mapped to potential symbols using
the dynamically configured thresholds to obtain a symbol
stream \( S \). Then the judging procedure **Algorithm 1** be-
gins to recognize valid symbols and discard illegal ones.
Finally, the valid symbols are mapped into bits to form
candidate packets, and upon receiving sufficient packets,
Raptor decoding procedure (omitted for brevity) is triggered
to recover original messages.

---

**Algorithm 1: Identifying Symbol in CASK.**

<table>
<thead>
<tr>
<th>Data: ( S, G'(i), G''(i) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Result: ( S_A )</td>
</tr>
<tr>
<td>begin</td>
</tr>
<tr>
<td>( S_A \leftarrow \emptyset );</td>
</tr>
<tr>
<td>while ( j &lt;</td>
</tr>
<tr>
<td>if ( G'(i) ) crosses zero in band ( j ) then</td>
</tr>
<tr>
<td>( S_A \leftarrow S_A \cup {S(j)} );</td>
</tr>
<tr>
<td>else</td>
</tr>
<tr>
<td>if ( G'(i) &gt; 0 ) and ( G''(i) ) crosses zero from</td>
</tr>
<tr>
<td>negative to positive then</td>
</tr>
<tr>
<td>( S_A \leftarrow S_A \cup {S(j)} );</td>
</tr>
<tr>
<td>else</td>
</tr>
<tr>
<td>if ( G'(i) &lt; 0 ) and ( G''(i) ) crosses zero from</td>
</tr>
<tr>
<td>positive to negative then</td>
</tr>
<tr>
<td>( S_A \leftarrow S_A \cup {S(j)} );</td>
</tr>
<tr>
<td>( j \leftarrow j + 1 );</td>
</tr>
<tr>
<td>end</td>
</tr>
</tbody>
</table>

---
3.5 Communication Capacity under CASK

Assuming the receiver can perfectly demodulate CASK symbols, now let us figure out the achievable data rate under high order modulation of CASK. While in CeilingTalk [29] a communication model was built only for OOK modulation, we hereby extend that communication model for higher order modulation, e.g., 4- and 8-ASK used in this paper. As previously analyzed in [29], the achievable data rate of a typical LED-Camera communication system is confined by the camera’s frame rate and the number of bits captured by a frame, and the number of OOK bits carried by one frame is bounded by:

\[ N_{\text{OOK}} \leq \left\lfloor \frac{L_{\text{Rol}} F_r}{W_r F_r} \right\rfloor, \]

where \( L_{\text{Rol}} \) is the length of the projection region of an LED transmitter working at a transmission frequency of \( F_r \), \( F_r \) is the camera receiver’s rolling-shutter frequency, and \( W_r \) is its corresponding width at given column resolution. We can readily re-use the same model for calculating the number of symbols captured by a frame, thanks to our CASK built upon purely digital control similar to OOK as mentioned in Section 3.2. Therefore, the number of symbols that a frame may contain should have the same upper-bound:

\[ N_{\text{symbol}} \leq \left\lfloor \frac{L_{\text{Rol}} F_r}{W_r F_r} \right\rfloor. \]

Therefore, the number of CASK bits carried by a frame is \( N_{\text{CASK}} = b \times N_{\text{symbol}} \), where \( b \) is the bit-per-symbol ratio\(^4\). The achievable raw bit rate of such an LED-Camera VLC system is:

\[ C = R N_{\text{CASK}} = b \times R \times N_{\text{symbol}} \leq b \times R \times \left\lfloor \frac{L_{\text{Rol}} F_r}{W_r F_r} \right\rfloor \]

where \( R \) is the frame rate of the camera receiver. Intuitively, the potential data rate could be improved by a higher order modulation because we can obtain a larger \( b \) comparing with the simple OOK.

4 System Evaluation

In this section, we evaluate the performance of proposed CASK modulation scheme with compositing light emissions and investigate achievable communication capacity under various ASK symbols and transmission frequency. We then discuss the experimental results with respect to various metrics.

4.1 Experiment Settings

We build the transmitter with commercial LED strips [44] and self-developed LED driver which can control each individual LED strips with low-cost transistors. The LED luminaire is made of 7 LED strips each carrying 36 LED chips, hence a size of 60 cm \( \times \) 7 cm similar to a common fluorescent luminaire, here we make a cover for the transmitter as common LED luminaires or it is too dazzling when we conduct experiments. Fig. 6 shows our testbed settings for field experiments. We slightly tune resistance of current-limiting resistor on the driver for each LED strip to maintain an appropriate brightness step thus a befitting grayscale variation in a frame.

In order to better evaluate the performance of various CASK, we fix the packet structure as a preamble of five successive brightest symbols under each order of CASK modulation, e.g. \( A_3 \) in 4-ASK and \( A_7 \) in 8-ASK, and a single darkest symbol \( A_0 \) to indicate the header, followed by 24 data bits (8-bit packet sequence number and 16-bit payload), and finally ends with a \( A_0 \). Under such a configuration, the transmitter sends identical payload bits but with different orders of CASK modulation, e.g. 2- 4- and 8-ASK. Fig. 7 graphically shows the benefit of receiving more packets in a frame under higher order CASK, while we will quantitatively evaluate this (termed as PFR) in following sections. We also consider OOK (2-ASK) as a baseline; it represents existing LED-Camera VLC systems (e.g., [11]). Raptor coding [45] is used as an FEC method to combat packet loss. Thus a message contains \( k \) packets, and Raptor coding encodes original \( k \) packets at the rate of \( n = 1.25 \times k \) to generate encoded packets. To simplify our evaluation, we fix \( k = 26 \) for throughput test and set the decoding overhead as 0.15.

A Nexus 6 smartphone is used as the receiver, and we build demodulation/decoding into an Android application on it. The decoding latency including Rol extraction and demodulation procedure (as described in Section 3.4) is around 30 ms, sufficient for real-time operation as it is less than the frame gap period of approximately 33.3 ms under a frame rate of 30 fps. Therefore, we configure the Nexus 6’s camera to work in the preview mode, and the frame rate is 30 fps. We fix the exposure time to 0.133 ms. Besides throughput, we also evaluate performance of various CASK modulation in

---

4. Here, \( b = \log_2(N + 1) \), where \( N \) is the number of grouped LED chips of an LED transmitter.
terms of Packet Frame Ratio (PFR) and Packet Error Rate (PER). The former is the ratio between the successfully identified packets (those between two consecutive headers) and totally used frame number, which makes sense as it directly indicates how many packets a frame can carry under various CASK modulation. The latter is the percentage of wrongly demodulated packets out of all successfully identified ones. Each experiment includes 50 sessions and every session contains 200 packets (before FEC); we report the average results over all sessions, except for throughput where peak values are reported as well.

4.2 ISO Impact on Demodulation

Apparently, a suitable ISO configuration for a camera receiver is crucial due to it directly impacts on band brightness (hence SNR). We first study the impact of camera setting on PFR and PER in this section. Here, we put the receiver at a distance of 0.4 m away the transmitter. Fig. 8 reports the experimental results. It is expectable that increasing ISO degrades both PFR and PER as it brings serious blooming effect resulting in boundary blurring between neighboring symbols. It is also obvious that 8-ASK is more sensitive to ISO setting comparing with the other lower order modulation, because symbol distance is shrinking with higher order CASK modulation. Nevertheless, CASK, e.g. 4- and 8-ASK, still manages a stable communication channel with an ISO configuration of below 200. Therefore, the ISO is by default configured at 200 for rest tests to maintain a reasonable trade-off between throughput and communication range. We leave it as future work integrating automatic ISO configuration into the demodulator.

4.3 Demodulating with Increasing Frequency

As analyzed in 3.4, a higher transmission frequency may yield a higher data rate, because a frame/RoI can carry more symbols (also packets). However, a higher frequency complicates demodulation due to it significantly shrinks the width of a band as shown in Figure 5. Therefore, we first evaluate demodulation performance under different frequency in terms of PFR and PER. We vary the frequency from 3kHz to 6kHz, yet put the receiver at a fixed position with a distance of 0.4m away the transmitter, and report the PFR and PER results with three different orders of CASK modulation in Fig. 9(a) and 9(b), respectively.

As we would expect, a higher order CASK yields higher PFR, because given identical transmitted payload (valid bits) the amount of mapped symbols is less. As a result, 4- and 8-ASK always have higher PFR than simple OOK as shown in Fig. 9(a). However, the PER is getting higher with an increasing frequency as shown in Fig. 9(b), since symbols are prone to be interfered by various noises such as blooming at a higher frequency resulting in thinner bands. It is quite straightforward that OOK achieves the lowest PER among three evaluated CASK modulation, yet it has to take less received packets (lower throughputs as evaluated in Section 4.6.1) as a price. In a word, high order CASKs, e.g. 4- and 8-ASK, always achieve higher PFR and maintain reasonable PER even at 6kHz compared to simple OOK, so it definitely shows the potential of boosting data rates via CASK. We hereafter set the transmission frequency at 5kHz.

4.4 Demodulation under Ambient Light

Since typical indoor environments have windows or non-VLC luminaires, the variance of ambient light may effect the performance of VLC and we hence evaluate the performance of CASK under varying ambient light in this section. We put the receiver at a distance of 0.4m away the transmitter, and put an extra non-VLC luminaire close to the transmitter to mimic the potential ambient interference. We use a light meter APP to monitor the illuminance and vary the illuminance from 1200lux (CASK transmitter only) to 1600 lux (sum of the CASK transmitter and the extra luminaire). Fig. 10 reports the experiment results, and we observe that the ambient light variance has almost no effect on PFR but some minor impact PER, especially for 4- and 8-ASK. Intuitively speaking, higher ambient illuminance does slightly reduce the SNR for each symbol (hence affecting PER), but not much to the packet header (indicating packets) due to its highest signal strength. Higher-order modulations are more prone to be affected by ambient light because of the
decreased symbol distance under such modulations. Fortunately, the ambient noise can be significantly suppressed by the extremely short exposure time in LED-Camera VLC (e.g., 0.133 ms in this paper), so the PER is still sufficiently low for our CASK, < 5% and < 10% for 4- and 8-ASK, respectively. Therefore, CASK for direct LED-Camera VLC is still a competitive choice even under strong interference caused by ambient light.

4.5 Channel Property

According to aforementioned analysis, various parameters, e.g., transmission frequency, impact on demodulation performance of CASK. We hereby investigate demodulation performance under different experimental settings.

4.5.1 Attenuation in Communication Range

We then evaluate the channel property with varying the transmission distance under three adopted CASK modulation. We change the distance from 0.2 m to 1.0 m due to the limited size of the used LED transmitter in our testbed, but it is quite easy to extend the transmission distance with a longer LED luminaire [11]. Again, as higher order modulation embeds more bits into a symbol leading to a shorter packet size in symbol level, 4- and 8-ASK double PFR comparing to OOK as illustrated in Fig.11(a). In particular, a frame hardly contains an entire packet with OOK at a distance of 1m, while it can still receive two packets averagely under 4- and/or 8-ASK. Such a higher PFR offered by high order CASK definitely implies its ability on boosting data rates. Moreover, a longer distance causes lower signal strength at the receiver side, hence reducing the dimension of effective RoI and also the grayscale, so Fig. 11(b) intuitively shows that both PFR and PER degrade with an increasing distance for all CASK modulation schemes. Since OOK has a larger symbols distance in grayscale (maximum vs. minimum), the communication distance has the least impact on its demodulation performance resulting in a stable PER. Nevertheless, CASK modulation still produces a reasonable PER of less than 2% under 4-ASK and 13% under 8-ASK within 1m.

4.5.2 Impact of Viewing Angle

We test the demodulation performance under various viewing angles for CASK modulations. We vary the receiver’s viewing angle within $[-60, 60]^{\circ}$ but maintaining the same distance of 0.4m from the LED transmitter’s center to emulate that a user may not face the transmitter perpendicularly. Fig. 12 reports the results in terms of PFR and PER, as we expected, the PFR is reduced with the increasing of viewing angle due to the deformed RoI that degrades the qualities of received symbols. To be more specific, OOK has a very stable channel, i.e. a lower PER, due to it is insensitive to change of grayscale, while 4- and 8-ASK both have their PERs increased suddenly when approaching $-60^{\circ}$ and $60^{\circ}$. Nevertheless, more packets are received under higher order CASKs in a frame (indicated by a higher PFR), so it may still offset the higher PER.

4.5.3 Impact of User Mobility

As a practical communication technique, it is imperative for LED-Camera VLC to provide service for mobile users, therefore, we hereby evaluate the demodulation performance of various CASK under user mobility. The basic experiment settings are the same as those stated in Section 4.1: each experiment contains 50 sessions and very session includes 200 packets. To better evaluate the impact of user mobility, this experiment is conducted by a user with three different motion patterns: i) Vertical Reciprocating Motion (VRM) that the user holding the smartphone moves back and forth when facing the transmitter, ii) Parallel Reciprocating Motion (PRM) that the user moves in parallel with the transmitter while keeping the smartphone facing the transmitter, and iii) Free Motion (FM) that the user freely wanders under the coverage (i.e., a half disk area with a radius of 1 m) of the transmitter.

Intuitively, lower order modulations, e.g., OOK and/or 4-ASK, have a more stable communication channel, i.e., a lower PER, as they have a larger symbol distance to combat the symbol distortion caused by image blurring under user mobility as shown in Fig. 13(b). Similar to previous evaluations, 8-ASK always has the largest PFR as shown in Fig. 13(a): it carries more packets in a frame thanks to its...
higher bit-symbol-ratio. Again, since OOK and 4-ASK have larger symbol distance comparing with 8-ASK, the former two lower order modulations achieve lower PER (below 15%) as shown in Fig. 13(b). Moreover, comparing with the results for static users, both PER and PFR degrade due to the channel impairments such as out-of-focus or even not capturing the transmitter when the user moves. Nevertheless, with the help of Raptor codes our CASK modulations could provide a reasonable throughput for mobile users as evaluated in following Section 4.6.4.

4.6 Throughput

In this section, we evaluate the throughput offered by various CASK modulation schemes. Whereas the above evaluations on channel property demonstrate possible communication ability with various CASKs, the throughput is evaluated from the perspective of realistic application on our testbed. The throughput in this paper is computed as the totally recovered data bits after Raptor decoding divided by the transmission time, here we fix the length of a message with 26 original packets (i.e. $k = 26$ as mentioned in Section 4.1). The transmission time is defined as the time span from starting receiving the first frame till the transmitted message gets decoded. We conduct experiments with OOK, 4- and 8-ASK based on our testbed, and report both maximum and average throughput for each experiment. In particular, we apply the complementary coding scheme proposed in Section 3.3 to 4- and 8-ASK, and employ Manchester coding for OOK to maintain DC balance so as to avoid flicker. We refrain from comparing CASK with CSK used in ColorBars [10], as it uses a Tri-RGB LED chip that is very hard (if not impossible) to scale up to full luminaires for a longer transmission distance, as explained in Section 2.4.

4.6.1 Throughput vs. Frequency

As studied in Section 3.4 and Section 4.3, transmission frequency impacts on the demodulation performance (hence the throughput). We hereby evaluate the throughput with various CASK modulation at a varying transmission frequency in this section. As shown in Fig. 14, both peak and average throughput get increasing with a higher frequency; it is totally agreed with the results of PFR evaluated in Section 4.3. 8-ASK always outperforms the other lower order modulation schemes below 5 kHz, while 4-ASK can support a higher transmission frequency up to 6 kHz as illustrated in both Fig. 14(a) and Fig. 14(b). The benefit of 4-ASK may stem from a higher SNR with grouped LED strips as described in Section 3.2 and relatively reasonable demodulation complexity. Nevertheless, the higher order modulation significantly boosts data rates comparing to the simple OOK, strongly demonstrating the effectiveness of our CASK.

4.6.2 Throughput vs. Distance

We then evaluate the throughput provided by CASK under a varying distance, and we report the outcome in Figure 15. As we expect, modulation schemes with higher order, e.g., 4- and 8-ASK, produce obviously higher data rates than OOK at any distance due to higher bit-per-symbol ratio. Apparently, 8-ASK achieves a higher peak throughput than 4-ASK, but a similar average throughput comparing to 4-ASK as illustrated in Fig. 15(a) and Fig. 15(b). The benefits of CASK over OOK become more evident at a longer distance, because the higher bit-per-symbol ratio of high order CASK shortens the packet length that in turn allows more packets in one RoI. In conclusion, our CASK can deliver an average throughput up to 2 kb/s at a distance of 1m with a single LED luminaire; it is almost fourfold throughput provided by conventional OOK [11]. If we further enhance the performance with collaborative transmission from multiple luminaires, the throughput should be high enough to support commercial applications, such as advertisement/coupon delivery in a shopping mall.

4.6.3 Throughput vs. Viewing Angle

In Section 4.5.2, we have evaluated the channel property in terms of PFR and PER under various viewing angles. We hereby present the achievable throughput by the proposed CASK modulations in this section. As reported in Fig. 16, higher order CASK modulations, i.e., 4- and 8-ASK, outperform the baseline of OOK in both peak and average throughput up to 8 kb/s at a viewing angle of 60 degrees.

![Fig. 14. Maximum and average throughput with a varying transmission frequency.](image)

![Fig. 15. Maximum and average throughput with a varying distance.](image)

![Fig. 16. Maximum and average throughput with a varying viewing angle.](image)
throughput thanks to their higher bit-symbol-ratios. The figures also show an evident impact of the viewing angle on the throughput, totally agreeing with the results presented in Section 4.5.2. In summary, high-order modulations can significantly improve the average throughput comparing with the simple OOK modulation even in the worst case scenarios, but their performances are more susceptible to an increasing viewing angle. In realistic application scenarios, users could often make a few more steps to reach the best view angle for a better communication service.

4.6.4 Throughput vs. User Mobility

We finally report the achievable data rate of various CASK for mobile users in this section. Under the same experiment settings as mentioned in Section 4.5.3, Fig. 17(a) and 17(b) graphically present the maximum and average data rates provided by our CASK modulations for mobile users, respectively. Comparing with the throughput by static users, the performance is indeed worse because user mobility significantly degrades the quality of captured frames by camera. For example, a slight joggle caused by user mobility could cause the camera out-of-focus thus distorting the shape of RoI and leading to incorrect demodulation, and the camera may even fail to capture the transmitter; these all result in a larger packet loss. Fortunately, advanced channel estimation and equalization algorithms (e.g., [14]) have potential to suppress such channel impairments to a large extent. Nonetheless, the experiment results shown in Fig 17(b) still demonstrate that our CASK modulations of 4- and 8-ASK significantly improve the throughput comparing with OOK and can offer a data rate of around 2 kbps even for mobile users.

5 Conclusion

In order to further boost data rates for LED-Camera VLC systems, this paper presents Composite Amplitude-Shift Keying (CASK) as a method to produce high order ASK symbols via compositing light emissions. Our novel idea relies on the nature of commercial LED luminaires consisting of multiple LED chips, so that CASK can control LED chips/stripes individually. As a result, light emissions get composed in the air so as to form different brightness levels (hence ASK symbols) on the receiver side. To handle the flicker issue caused CASK modulation (given that existing RLL codes are only devised for low order modulation, e.g. OOK), we have proposed a complementary coding mechanism to maintain DC balance at both symbol and individual LED chip levels for CASK modulation. Moreover, we have engineered a delicate demodulation algorithm combining conventional width-based demodulation with the gradient of grayscale to effectively demodulate CASK symbols. We have built a prototype for CASK and performed extensive field experiments based on it; the results have demonstrated that CASK can be used in LED-Camera VLC systems with commercial LED luminaires and achieve a throughput of up to 2 kbps at a 1 m distance with only a single LED luminaire, almost fourfold of the throughput offered by conventional OOK modulation. We believe that applying CASK to LED-Camera VLC would be able to push the COTS-enabled VLC towards practical deployments for realistic applications in need for reasonable data rates.

In terms of future work, we plan, on one hand, to extend the scale of our prototype for conducting more comprehensive evaluations on both higher-order modulations (beyond 8-ASK), as well as on collaborative transmissions among neighboring LED luminaires. On the other hand, we are on the way to introduce advanced channel estimation algorithms into the demodulator, as taking channel effect into account may further improve the demodulation performance and in turn yield a higher throughput. Such a scale-up prototype could help us to identify more solutions to further improve the performance for LED-Camera VLC.

Acknowledgement

The authors would like to thank the anonymous reviewers for their constructive feedback and valuable input. This work was supported in part by the Fundamental Research Funds for the central Universities under Grant No. 20822041B4232, National Natural Science Foundation of China under Grant No. 61373091, Sichuan Science and Technology Program under Grant No. 19ZDYF0045 and 19CXTD0005, the AcRF Tier 2 Grant MOE2016-T2-2-022, and the DSAIR Center at NTU.

References


Yanbing Yang received his B.E. and M.E. degrees from University of Electronic Science and Technology of China, China, and his Ph.D degree in Computer Science and Engineering from Nanyang Technological University, Singapore. He is currently an Associate Research Professor in College of Computer Science, Sichuan University, China. His research interests include IoT, visible light communication, visible light sensing, as well as their applications.

Jun Luo received his BS and MS degrees in Electrical Engineering from Tsinghua University, China, and the Ph.D degree in Computer Science from EPFL (Swiss Federal Institute of Technology in Lausanne), Lausanne, Switzerland. From 2006 to 2008, he has worked as a post-doctoral research fellow in the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada. In 2008, he joined the faculty of the School Of Computer Science and Engineering, Nanyang Technological University in Singapore, where he is currently an Associate Professor. His research interests include mobile and pervasive computing, wireless networking, applied operations research, as well as network security. More information can be found at http://www.ntu.edu.sg/home/junluo.