System Design of LED-to-Rolling-Shutter-Camera Communication using Color Shift Keying.

Merdan Durmus[*]
Supervisor(s): Koen Langendoen, Marco Zuñiga Zamalloa, Miguel Chavez Tapia[†]
EEMCS, Delft University of Technology, The Netherlands
20-6-2022

A Dissertation Submitted to EEMCS faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering

Abstract

Visible Light Communication (VLC) is becoming an important research area where Visible Light Sources such as LED, Halogen Lamps and even the sun can be used for Wireless Communication. LED-to-Camera Communication is a form of VLC, where the camera is able to notice intermittent stimuli above a certain threshold frequency (*flicker fusion frequency*) which human eyes cannot notice. There exist numerous modulation methods for LED-to-Camera Communication, however with some modulation methods VLC cannot support any practical applications because the achievable data rate either is too low or the system is difficult to be implemented on off-the-shelf devices.

In this paper, we will analyse the data rate of LED-to-Camera Communication with Color Shift Keying (CSK) as a modulation method to encode the transmitted data, furthermore this paper provides a design for a system which can obtain a more practical data rate by using CSK as modulation method and with a CMOS image sensor as receiver. A CMOS image sensor consists of a matrix of photodiodes which combined make up the captured image. The rolling shutter effect makes use of this matrix like structure by scanning each line separately. This paper also provides hypothesis insights into the throughput and goodput of different CSK-Modulations. From the hypothesis follows that a lower CSK-Modulation method can achieve a much higher goodput. The low symbol error rate of lower CSK-Modulations makes it an excellent choice for reliable Tri-LED-to-Camera Communication.

1 Introduction

All wireless communication is done using radio waves, and it is becoming crowded. Imagine seeing all of these radio waves around you, we would see Cellular, WiFi, AM/FM, BLE, etc. In the last couple of years, researchers have started to investigate the use of light to create wireless links to suppress the demand on Radio Frequency bands [Hu et al., 2015; Kuo et al., 2014; Zhice Yang, 2015; Zhao Tian, 2020].

In VLC, LEDs can transmit information using different modulation techniques and can be received by either a photodiode or image sensor (camera). This kind of communication is important because it enables any mobile device equipped with a camera to receive data from a LED. A great advantage of LED-based VLC is that it can offer specific information. Since the LED must be visually located by the mobile device user, the LED can offer specific information for that transmitter.

There already exist numerous applications which make use of visible light communication, these applications include designs of a new generation of toys [N.O. Tippenhauer and Mangold, 2012] and accurate indoor positioning systems [Yu-Lin Wei and Lin, 2017; Julian Randall and Burri, 2007] to human sensing [Tianxing Li and Zhou, 2015]. The above applications are transforming light sources to actively modulate data, this implies turning on and off LEDs. Since human eyes are sensitive to low-rate changes in light intensity, the transmitter must transmit pulses at a rate over 80 Hz [Michal, 2021] to avoid flickering. Most cameras' sampling capabilities are 240fps fps, which is far below the high pulse rate. The cameras' sampling capabilities do ensure a flicker-free LED-to-Camera Communication, but to achieve a high data rate we must have a high pulse rate.

To handle this problem researchers have found a new camera sampling technique, leveraging the rolling shutter effect of the CMOS image sensor to decode high-rate pulses from the high-resolution image. A CMOS image sensor consists of a matrix of photodiodes which combined make up the captured image. The rolling shutter effect makes use of this matrix like structure by scanning each line separately. However, the rolling shutter effect used by the receiver imposes some limits on the type of modulation that can be used by the transmitting LED. Furthermore, since the CMOS image sensor needs time to process the captured frame before it can move to the next frame, there exists a gap in-between the frames. During this inter-frame gap the transmitted symbols are lost.

Furthermore, a bigger limitation of LED-based VLC is the fact that LED-to-camera communication has a low achievable data rate. Such a low data rate would limit the transmitter to only transmit simple identification codes etc., it is unfortunately not sufficient for practical applications which require the transmitter to transmit small images and textual content

In this paper, we will design, and analyze the data rate of, an LED-to-Camera Communication system which uses Color-Shift Keying (CSK) for modulating the data emitted by the light source (LED). This paper will also provide insights into different Color Shift Keying schemes.

The system makes use of a Tri-LED which can emit Red, Green and Blue (RGB) values to produce white light. With the use of Tri-LED, any color can be created by combining the RGB intensities. By using CSK we reduce the symbol duration significantly compared to other modulation schemes, such as Frequency Shift Keying (FSK) based modulation schemes. FSK makes use of different frequencies at which it transmits long runs of ONs and OFFs. The reduction in symbol duration is because with CSK, the transmitter is able to send more bits associated to a color, in other modulation schemes the transmitted color or frequency corresponds to a single bit. This reduction in symbol duration and the higher color modulation means an improved data rate compared to other systems [Hu et al., 2015]. The design of this system does have to take the following research challenges into account:

- LEDs' main purpose is illumination, the system should make sure that the color-modulation does not affect the human-perceivable color.
- Rolling shutter cameras suffer from an inter-frame gap where the transmitted symbols during the gap are lost.
- The current cameras in mobile devices are very diverse, they may interpret the transmitted color symbol differently due to differences in the type of color filter in front

of the photodiodes in the CMOS image sensor.

We propose a system design for LED-to-Camera Communication which can be build using off-the-shelf hardware. The design challenges found during the system design are explored and solutions are proposed. Furthermore, the used modulation scheme, Color Shift Keying, is analysed and hypothesis are formed. Below we summarize the main contributions of this paper:

- We propose a system design for LED-to-camera communication with off-the-shelf hardware components.
- We explore the design challenges that arise when using Color Shift Keying as a modulation method
- We explore the design challenges that arise when using a CMOS image sensor as a receiver.
- We analyse a color-based modulation scheme which takes the diversity of the cameras and the inter-frame data loss into account.

2 Background

In this section, we explain how the human eye perceives stimuli and explain how we can change stimuli without humans noticing, then we dive into the background of camera sensor's rolling shutter effect, after which we will discuss different modulation techniques such as On-Off keying and Frequency Shift Keying and most importantly Color Shift Keying.

2.1 Human Perceivable Stimuli Changes

The human brain needs a certain amount of time to 'process' the received images by our eyes, our visual system cannot respond to immediate changes in stimuli. This gives a delay in observing changes in luminance or color. This gives that our visual system can only perceive intermittent stimuli below a specific rate. However, above that rate the human brain cannot cope with changes in luminance or color. This threshold is revered as the flicker fusion threshold [Landis, 1954].

By using the flicker fusion threshold, we can generate different colors with the Tri-LED at a high frequency which is not perceivable by humans. The combined color perceived by the human eye will be white, since white is the presence of all colors generated by the transmitter. However, the CMOS image sensor will notice the change in colors.

This approach is easy and effective but it can still lead to human perceivable stimuli changes. A long run of the same symbol will cause the transmitter to generate the same color for a while. This may be perceivable for human eyes.

2.2 Rolling Shutter

The most commonly used image sensor (camera) in mobile devices nowadays such as smartphones, tablets and laptops is a Complementary Metal-Oxide Semiconductor (CMOS) image sensor. The CMOS image sensor consists of a matrix of photodiodes which scans one line at a time rather than scanning a scene all at once. Each photodiode in the image sensor converts the incoming photons to voltage, where the voltage is used to obtain the pixel value. This technique of scanning photodiodes one line at a time is referred to as the rolling shutter effect. Rolling shutter increases the sampling rate by up to thousands of times per second. The final image produced by the image sensor is a concatenation of all the lines of the photodiode. Figure 1 shows how the rolling shutter effect can be used to alternate between ON and OFF states and how the camera produces an image. However, the usefulness of the rolling shutter effect is limited for two reasons:

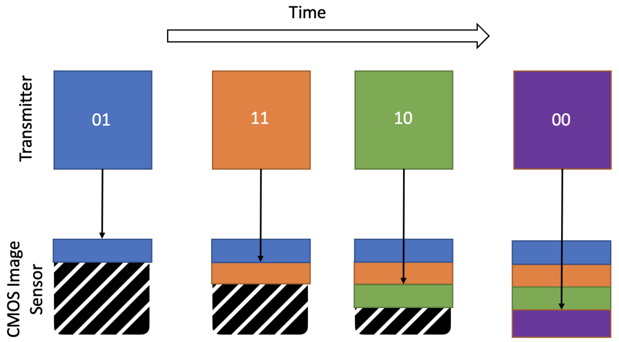

Figure 1: Rolling Shutter Effect of the CMOS Image Sensor

1. The camera must have a very high resolution if it is not close to the transmitter. This is because the transmitter is only a small part of the final image. If there are not enough pixels used then it is impossible to carry the message.
2. Processing high-resolution images can cause a very large computational overhead at the receiver.

2.3 Modulation Techniques

Thanks to the rolling shutter effect, the communication between a camera and LED is achievable where we can have multiple data symbols within one camera frame. The modulation of these multiple data symbols can be achieved using different modulation techniques. Here we will discuss three of those modulation techniques, we will introduce On-Off Keying, Frequency Shift Keying and Color Shift Keying.

On-Off Keying (OOK):
In OOK modulation the two different states of the LED, ON and OFF, are used in the communication. The two different states with regard to the rolling shutter effect can be seen in figure 1. This modulation method is, unfortunately, less robust against ambient light since it only utilized the white light emitted by the LED. Another limitation is the amount of ON and OFF states transmitted by the LED after each other. If the LED transmits a long run of ONs (1s) or OFFs (0s), this will lead to flickering which can be seen by the human eye.

Frequency Shift Keying (FSK):
FSK makes use of different frequencies at which it transmits long runs of ON and OFFs as seen in figure 2. Due to longer symbol duration of FSK, it reduces the demodulation error. FSK also introduces multiple ON-OFF bands in each symbol.

Color Shift Keying (CSK):
Color shift keying is a modulation method which exploits the design of many commercial LED luminaires which make use

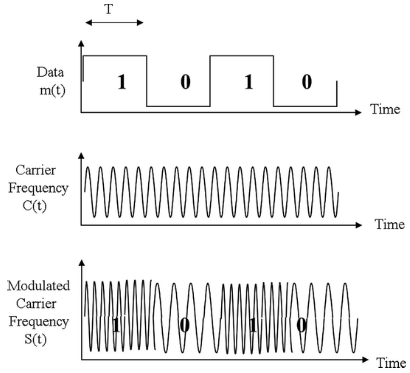

Figure 2: FSK waveforms. The frequency of the carrier changes in accordance with the input digital signal. The amplitude of the carrier remains the same, from Faruque, S. (2017). Frequency Shift Keying (FSK) [Faruque, 2017]

of three separate LEDs (Red, Green, and Blue) to generate light. By making use of different color patterns to modulate data, the achievable throughput is much higher compared to OOK and FSK. CSK makes use of the three different LEDs to generate a variety of colors, by changing the intensities of the three colors. Since there is a huge diversity in mobile devices, each with a different RGB color space, we need a standard color space to unify color standards. This color space is defined by CIE 1931. This color space is used by CSK to form a constellation triangle inside the color space. After forming the constellation triangle, the constellation points can be chosen inside the constellation triangle. The number of constellation points is dependent on the degree of CSK. The degree of CSK means how many symbols will correspond to a color. However, when choosing these constellation points, their inter-symbol distance should be maximized to decrease the inter-symbol interference. Given the RBG coordinates of the constellation triangle, the CSK constellation points can be calculated from the specification in [Yoon et al., 2019].

The intensities of the three different LEDs (Red, Green, and Blue) are controlled with a technique called Pulse Width Modulation (PWM). PWM uses digital signals to control the power supplied to the LEDs. This digital signal of the PWM can switch between full voltage and no voltage, which represents on and off respectively. The percentage of the time the digital signal is on over a period of time is described by the percentage duty cycle. Different duty cycles in an LED means that we can obtain a steady voltage between 0 and full, this allows us to control the brightness of the LED. By using three different PWM signals with different duty cycles for each LED (Red, Green, and Blue) in a Tri-LED, the Tri-LED will generate an accumulated color.

This accumulated color is then received by the camera which maps the received colors to the reference colors for demodulation. When using a higher degree CSK, the receiver should also be able to support the computational overhead it causes at the receiver side, together with the image sensor which must support all these colors. However, when using a higher degree of CSK, more inter-symbol interference can be caused because more colors means that it becomes more difficult for the receiver to distinguish between different colors. The major challenges involved in using CSK with rolling shutter cameras will be identified and addressed in this paper.

3 Related Work

LED-to-camera communication has been comprehensively researched in the last decade. [Tsai et al., 2014; Rajagopal et al., 2014; Roberts, 2013; Danakis et al., 2012]. Different modulation schemes have been researched as in [Roberts, 2013] where R.D. Roberts proposed the use of undersampled OOK. In undersampled OOK each frame only contains a single symbol in order to reliably deliver data even in highly noisy environments. The combination of the rolling shutter effect and OOK modulation has been studied in [Danakis et al., 2012] where C. Danakis et al. proposed a shorter symbol duration. However, OOK suffers from severe noise due to ambient light. Together with the fact that undersampled OOK can result in flickering effect due to the very slow symbol rate. In [Tsai et al., 2014] H.-Y. Lee et al. proposed a system design with FSK modulation to deal with the synchronization challenge and the inter-frame gap loss. Unlike these studies our system design will be modulated with color-shift keying

CSK has been studied very well in the past [Zhice Yang, 2015; Chan et al., 2017]. These studies make use of passive VLC where they don't control the light source. They introduce the concept of liquid-crystal shutters and polarizators together with dispersors to introduce color patterns. The color patterns are then modulated using binary CSK (BCSK) Unlike these studies, our system design assumes we have full control of the transmitting LED. Other papers have studies CSK as well [Monteiro and Hranilovic, 2014; Monteiro and Hranilovic, 2012a], these studies, unlike this paper, have not used a rolling shutter camera as a receiver.

Rolling shutter cameras have also been studied in visible light indoor localization [Kuo et al., 2014; Zhice Yang, 2015]. In these studies a rolling shutter camera is used that can receive bacons from multiple LEDs indoor and localize the mobile device using the angle of arrival calculation. Different from these, the main focus of this paper is to provide a high data rate and reliable communication in LED-to-camera communication instead of indoor localization. [Tsai et al., 2014]

LED-to-camera communication has been studied well in [Hu et al., 2015]. P. Hu et al. have found that their system ColorBars can achieve a data rate of 5.2 Kbps and 2,5 Kbps on a Nexus 5 and iPhone 5S respectively. In this paper we will replicate their approach on the LED-to-Camera Communication System and use a more high-end mobile device to evaluate and analyze the data rate of LED-to-Rolling-Shutter-Camera Communication with Color Shift Keying as modulation scheme.

4 Responsible Research

In this section, the ethical side of the experiments are discussed. First, we will discuss how we have handled the data during the experiments. After which we will explain the ethos of science. Finally, the health risks of LED flickering are explained.

4.1 Data Handling

During the experiment phase of the research, data points have been found regarding the performance of the LED-to-Camera Communication system. These data points have been used to evaluate the LED-to-Camera Communication system. It is important to note that all these data points were real and have not been made up. This research did not fabricate data points. The found data points have not been cherry-picked for their 'usefulness', thus this research did not do any data trimming. It is also important to note that the experimental setup described in the paper, is the same experiment setup used. During the experiment, there was no manipulation of the research materials used.

4.2 Ethos of Science (CUDOS)

The ethos of science can be described by the CUDOS principle. CUDOS stands for Communism, Universalism, Disinterestedness and Organized Skepticism.

Communism:
Communism means that research results should be open and accessible for all. Research builds up on each other, therefore it is of uttermost importance that research results are publicly available. The research results found in this paper are therefore all public.

Universalism:
The term universalism means that the evaluation of research results should not include prejudice against sexual orientation, gender, or even a person's scientific reputation. In this research, this is not applicable, except for the fact that our research keeps in mind that LED flickering can have a negative health impact on persons with autism. Our research does not include any form of prejudice.

Disinterestedness:
It is of uttermost importance that researchers are emotionally detached from their field of study and that they are pursuing the truth. Their opinions should not affect their research. This research has been done with a completely open mind. No political, economic, or religious interests have played a role during the research process.

Organized Skepticism:
During the research process, the involved parties have been critical of this work. This research has also been peer-reviewed by other students and responsible professors. In this paper, we have tried to be as open as possible about possible weaknesses.

4.3 Health Risks LED Flickering

During LED-to-Camera Communication, the communication link between the LED and Camera is established using LED Flickering. However, LED flickering can be harmful to human health when the flickering is visible to the human eye. Some effects of flickering include seizure, stroboscopic effects, migraine, exacerbation of repetitive behaviour in persons with autism, and asthenopic effects including eyestrain, fatigue, and reduced performance on visual tasks. Because these effects need to be taken seriously, we decided that we need to work with LED flickering above the flicker fusion threshold. The flicker fusion threshold is the threshold where human eyes are not able to keep up with the speed of changes in illumination. So by working above the flicker fusion threshold, we ensure that human health is not affected.

Figure 3: All components of the experiment, Redmi Note 10 as receiver, Tri-LED with Arduino and Diffuser as transmitter and a Tripod to hold the smartphone still

5 Methodology

The goal of this experiment is to create a link between an LED and Camera using the rolling shutter effect of the receiving camera. To overcome the limitations of low achievable data rate, the system makes use of CSK to modulate data. The Tri-LED transmitter is implemented using an Arduino Due board to control the intensities of the individual LED components in the Tri-LED. The full transmitter (shown in figure 3) consist of an Breadboard 400 points equiped with a LED (SML-LX1610RGBW/A), Potentiometers (RV24AF-40-15R1-B1k-3LA), and SMD resistors together with two Pin Headers. As a receiver, a Redmi Note 10 smartphone was used. Note that the Redmi Note 10 has a rolling shutter camera with a resolution of 4K/3840x2160 at 30fps, 1080p/1920x1080 at 60fps and 1080p/1920x1080 at 30fps. Furthermore, a tripod was used to hold the phone still during the data transmission. This is necessary since even a slight deviation of the position can increase the symbol error rate. Since the transmitter consists of a Tri-LED, which is a package of three LEDs, Red/Green/Blue, we need a diffuser (figure 3) to evenly distribute the incoming light. If we do not use the diffuser, the captured image by the camera will display some areas which are more red, green or blue, depending on where the individual LEDs lie underneath the cap of the tri-LED. Finally, since ambient light has an negative influence on the symbol error rate, this paper chose to do the experiments in the dark to minimize ambient light impact. All the components of the system can be seen in figure 3.

The captures images are then processed offline. This paper will evaluate the transmitter and receiver using three performance metrics: Symbol error rate, throughput and goodput.

Color Space Conversion: By converting the received frame from RGB color space to CIELab space, we remove the majority of the effect of brightness variation within the same color symbol. We find the first big change in color difference in the center column of the received frame and calculate how wide each stripe roughly should be. Then we can access the pixel values in the center column in the middle of each individual stripe. By using CIELab space we observe much smaller variance due to the removal of most of the brightness effects compared to RBG space.

Symbol Error Rate: By looking at the Symbol Error Rate (SER) we can see the demodulation errors experienced by the receiver due to inter-symbol interference. We measure the SER in CIELab space for the Redmi Note 10, these measurements are done using different modulation schemes. First the SER is measured for 8-CSK, then for 16-CSK and finally for 32-CSK. Since we also need to capture the effect of symbol frequency (the number of symbols per second), we vary the transmitter to transmit symbols from 1000 Hz to 4000 Hz, we increment it in steps of 1000 Hz for each modulation scheme. Note that the ISO settings and exposure time of the receiver is not modified.

Throughput and Goodput: After we have computed the Symbol Error Rate, we can compute the Throughput and the Goodput of our system in a similar way to the SER results. During the computation we again vary the modulation scheme and symbol frequency to investigate their impact on throughput and goodput. Note that we again don't modify the ISO settings and exposure time of the receiver. To calculate the raw throughput, we do not perform any error correction at the receiver side, we simply measure the number of received symbols. Now that we have computed the raw throughput, we can calculate the goodput of the mobile devices under different modulation schemes and symbol frequencies. The goodput is measured after RS error correction is done by the receiver. Now we only measure the correctly received or recovered symbols.

6 Experimental Setup and Results

6.1 Avoiding Human Perceivable Color Flickering

The main purpose of the transmitting LED is illumination, thus the system design must make sure that the human eye is not able to notice the transmitted colors. The human perceivable color of the transmitter must be white (unless otherwise specified). This can be achieved in multiple ways, the system can opt to introduce white light symbols without any meaning like in [Hu et al., 2015] or chose to use a constellation triangle such that the average color is white as defined in [Monteiro and Hranilovic, 2012b] or the system can even opt to use a slightly different modulation scheme, Color Sequence Shift Keying (CSSK) to avoid human perceivable color flickering.

Introduce White Light Symbols

The approach of introducing white light symbols is easy and straightforward. When the packetization of the data stream is finished, the system will insert white light symbols without meaning to keep the overall color white. This percentage of white light symbols introduced depends on the sent packet and the average color of the packet. When a sufficient number of these white light symbols are introduced, the system can guarantee a white light for human eyes. However, since these white light symbols serve no data transmission purpose, the system must reduce the number of white symbols used. In the paper, [Hu et al., 2015], the authors experimented with how many of these white light symbols should be used to still guarantee white perceivable light by human eyes. The authors increased the symbols frequency from 500Hz to 5000Hz where each symbol was randomly chosen from the constellation triangle. The percentage of white light symbols was varied and ten volunteers were asked to observe the LED light for color flickering. Figure 4 was found from their experiments, it shows the minimum percentage of white light symbols necessary to avoid color flickering.

[Hu et al., 2015] also found from this experiment that as the symbol frequencies increase, the number of white light symbols necessary to guarantee white light decrease. This follows from the fact that at a high symbol frequency, the symbols are more likely to be evenly distributed within the constellation triangle, which results in a white perceivable light by humans.

From the experiments of [Hu et al., 2015] follows that a transmitting LED at a high symbol frequency can achieve a higher data rate than at a low symbol frequency and at a high frequency the system needs less white light symbols to guarantee white perceivable light. The decrease in white light symbols, will again increase the data rate of the system since the system can send more data symbols then a low symbol frequency system during the same time frame.

A high symbol frequency system can achieve a much higher data rate then a low symbol frequency system, however, a high symbol frequency system does impose some challanges. The hardware used for the transmitter can limit the high symbol frequency, when an LED with a low maximum frequency is used, the system cannot use a higher symbol frequency. Furthermore, there are also some limitations at the receiver side. The processing of a high symbol frequency system becomes harder than a low symbol frequency system, When a high symbol frequency system is used, the captured image by the receiving rolling shutter camera has smaller color bands than a low symbol frequency system. This decrease in the width of the color bands makes the demodulation difficult and may introduce errors during demodulation.

Average Constallation Triangle

NOTE: from paper: *”Constellation design for color-shift keying using interior point methods”*, but I don't think this applies for this paper, TODO: read again and evaluate

Color Sequence Shift Keying (CSSK)

The idea behind Color Sequency Shift Keying is similar to CSK, but with a main difference. Instead of using each color as a symbol, CSSK will use a sequence of colors as a symbol. With CSK it is difficult to achieve a balance between Red, Green and Blue because the balance is dependent on which symbols are transmitted in the package. This balance is called the color balance and it is the vector representing the average perceived color transmitted by the LED. In [Monteiro and Hranilovic, 2012b], the authors explain how the color balance can be calculated for a packet. However, since in CSSK each symbol is a sequence of colors, the color balance is easier to achieve. When using CSSK, the system must also use a delimiter which is introduced in between two sequences. Since our rolling shutter camera suffers from an inter-frame data loss, it is possible that the transmitted symbols are lost or partially received. The use of the delimiter enables the transmitter and receiver to synchronize with each other.

The main difference between CSK and CSSK is the number of bits that are represented by a color(sequence). This makes that CSSK represents less bits compared by CSK. If

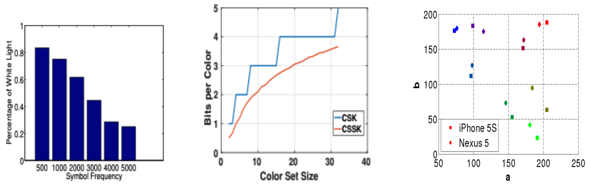

Figure 4: Percentage of white light symbols necessary to prevent color flickering, Efficiency of CSK and CSSK and Comparison on how the Nexus 5 and iPhone 5S perceive the same color, from Hu, P. (2015). ColorBars [Hu et al., 2015]

the size of symbol space is N! in CSSK, each symbol could represent $\log_2(N!)$ bits. Therefore, the number of bits *n* each color could represent is:

$$n = \frac{\log_2 N!}{N}$$

Whereas, for CSK, each color could represent

$$n' = \log_2 N$$

From this follows that $n' > n$, therefore, the efficiency of the CSK scheme is better than CSSK which can be seen in figure 4.

Even though, CSSK represents less bits than CSK for the same color, it does also have benefits over CSK. When using CSSK, there is no need for the introduction of white light symbols, which will increase the data rate for CSSK. Furthermore, CSSK has the property that it can achieve white light balance independent on the color rate. So even though the symbol length is longer in CSSK compared to CSK, the ability to achieve white balance without introducing white light symbols makes the two modulation schemes comparable.

6.2 Inter-Frame Data Loss

When using the CMOS image sensor as a receiver, the system suffers from an inter-frame data loss. This is because the captured frame by the receiving camera needs a certain amount of time to be processed. During this time frame, information transmitted by the LED is lost, this lost data stream is refered as the inter-frame data loss. There are multiple ways to retrieve the lost data during these consecutive frames. This paper will focus on error correction coding and packetization to handle the problem of inter-frame data loss.

Reed-Solomon Coding (RS Coding) Reed-Solomon Codes are block-based error correction codes, with error correction codes the system can retrieve the inter-frame data loss. The system must first divide the data stream into blocks of *n* bits. The inter-frame data loss can occur anywhere in the block of *n* bits. In $RS(m, n)$ coding, a codeword of *m* bits is generated by adding $m - n$ parity bits to the *n* data bits. In RS encoding, the system can detect errors up to $2t$ bits and correct up to t bits where $2t = m - n$. The detection and correction of bit errors can happen anywhere in the codeword of m bits, which is suitable for our system since the bit errors can happen anywhere in the block of *n* bits.

The size of the parameters n and m depend on the inter-frame gap of the receiving CMOS image sensor. However, the system design has to take into account that as the size of n increases, the computational overhead at the receiver increases which is not desirable for mobile devices.

To compute the inter-frame gap, the inter-frame loss ratio must be computed first. The inter-frame loss ratio is the ratio of size of inter-frame gap to the total size of a frame and an inter-frame gap. If we define S as the symbol rate which defined the symbols per second, the inter-frame loss ratio as l and F as the frame rate, we find that the number of symbols in one frame is defined as $F_s = (1 - l) * \frac{S}{F}$. However, as discussed in section 6.1, the system also includes white light symbols or uses CSSK as modulation scheme. μ_S is the illumination ratio in which the white light symbols are added, this ratio defines the ratio of the number of useful data symbols to the total number of data and white light symbols. If we define C as the size of bits in a symbol, the number of symbols lost between two consecutive frames is $L_S = l * \frac{S}{F}$, and the total number of data bits lost is $\mu_S * C * L_S$. The size of a codeword is then calculated as $m = \mu_S * C * (F_S + L_S)$, and in order to recover these lost bits, the RS encoding should define $t = \mu_S * C * L_S$, which means that the parity size is $2 * t = 2 * \mu_S * C * L_S$. Therefore, the system must define $n = m - 2 * t = \mu_S * C * (F_S - L_S)$.

If the system choses to opt for CSK with white light symbols, the bitstream is first mapped to symbols based on the chosen CSK modulatuion scheme, and then the white light symbols are added afterwards. However, if the system choses to use CSSK as its modulation scheme, the bitstream is directly mapped to the symbols.

6.3 Receiving Camera Diversity

There exist many rolling shutter cameras on the market, each one is different from each other. This diversity of the receiving camera makes the communication between LED and camera extra difficult since the camera image sensor used in smartphones may differ in how they capture different colors. There are two main problem involved in the use of rolling shutter cameras. The first problem is that the receiver can be diverse in terms of how they perceive the transmitted color. To overcome this problem, the system makes use of calibration packets, which will be explained later in this section. The second problem with rolling shutter cameras is that they can differ in inter-frame loss ration, making the encoding differ for different mobile devices. This second problem is more of a compatibility problem and it should depend on the system designer to include a certain range of devices with a inter-frame loss ratio in the supported inter-frame loss ratio interval of the system.

Calibration packets

An CMOS image sensor consists of a matrix of photodiodes. These photodiodes capture the intensity of the incoming light and not the color. In order to capture the colors in the image, each photodiode is covered with a color filter. The Bayer filter is one of many filters used for this purpose. A Bayer filter is a matrix of filters with alternating rows of green-blue and green-red filters. This high number of green filters is explainable because the human eye is more sensative for the color green. Figure 4 shows that different cameras can perceive the same color differently. This is due to the different color filters used in the camera sensors.

Furthermore, the same camera may interpret the same symbol transmitted at a different time differently. This is because an captured image is controlled by many different parameters of the camera sensor (such as the exposure time, ISO, exposure value (EV) etc.). Many modern day cameras adjust these parameters automatically depending on the ambient light condition. The exposure time is the time that the camera shutter is open to allow the light into the photodiode matrix. A larger exposure time means that each photodiode will have more time to accumulate the incoming photons. The ISO is a parameter that determines how many photons are enough to saturate the photodiode. A lower ISO value means less sensitivity to light, while a higher ISO means more sensitivity. Exposure Value (EV) is simply a way to combine shutter speed and aperture to a single value. Although shutter speed an aperture both carry a lot of "side effects" like motion blur and depth of field, EV doesn't take those into account. EV only relates to exposure.

To tackle these problem in which the receiver may interpret the same color as different or where receivers differ in how they interpret colors, the system makes use of Calibration Packets. These calibration packets are send our by the transmitter. A calibration packet contains a preamble and all symbols of the modulation scheme. The preamble should be an easy to recognize pattern such as white, black, white, black followed by all the symbols of the modulation scheme. Once the receiver recognizes this pattern it stores the different colors corresponding to each symbol and uses the stored values for future demodulation. Because these calibration packets are sent out periodically, new receivers are also able to receive data after they have receiver a first calibration packet. Even channel conditions (such as ambient lightning and different parameters for the camera) are taken into account by using these calibration packets, because when the channel conditions change, the colors in the calibration packets change with it as well.

6.4 Demodulation Method

A rolling shutter camera captures a video recording of the transmitter and then extracts the symbols in each frame. Note than when using a color space for the transmitter, it must be chosen such that the symbols are equally distributed among all RGB wavelengths to product human perceivable white light. However, the receiver can after capturing the image, use any color space it desires. An easy way to do this, is to use the RGB color space and apply a distance metric between the saved colors in the memory of the receiver (from the calibration packets) and the received color. The receiver will then match the received color to the closest symbol color in its memory and by doing so the receiver can demodulate the received color. However, the use of CIELab color space is better suited for the demodulation as it can distill symbol's color by removing most of the effects of the brightness. This is necessary because for some receiving cameras, the edges are darker then the centre of the captured image as can be seen in figure 5. CIELab is a color space with three channels like RGB, the main difference is that CIELab has one channel for the lighness L and two channels for colors, a and b. The CIELab color space is shown in figure 5. From this figure follows that the a dimension spans from green (a) to red (+a) and the b dimension spans from blue (b) to yellow (b). Together with the vertical axis L which captures the brightness from black to white, every color can be respresented as $\{L, a, b\}$, however after removing the brightness dimension which the system doesn't need for CSK, we can respresent any color as $\{a, b\}$.

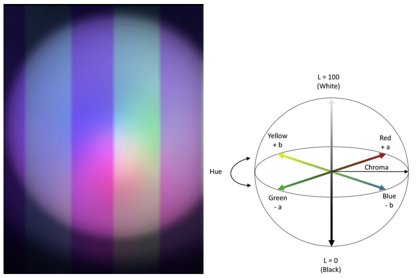

Figure 5: Single frame from the iPhone 12 and CIELab Color Space, from Ly. (2020). [Ly et al., 2020]

The demodulation process consists out of three steps which are explored in detail below:

Convert to CIELab As explained earlier, the light distribution in one frame is not evenly distributed as seen in figure 5. In this figure it is easy to see that the edges of the frame are darker than the centre which is brighter. By converting the received image to CIELab space, the system is able to eliminate the brightness dimension and still use the color by representing any color as $\{a, b\}$.

Finding the color bands After converting the received image to CIELab space, the system must find the different color bands. Then the color values are extracted from the centre of each band to reduce inter-symbol interference.

Symbol Matching Now that the system has found the different color bands, and has extracted the symbol colors from the calibration packet, it can find the corresponding symbol to each color band by calculating the Euclidean Distance between each color band and pre-saved symbol color. The Euclidean distance between two colors in CIElab space is calculated as follows:

$$\Delta E = \sqrt{(P_a[i] - R_a[j])^2 + (P_b[i] - R_b[j])^2}$$

, where $P_a[i]$, $P_b[i]$ are the two dimensions *a* and *b* of CIELab space of the color of the centre of a color band P[i], while $R_a[i]$, $R_b[i]$ are the two dimensions a and b of CIELab space of the color of the reference color symbol. We know from [Sharma, 2002] that the difference between two colors is noticeable when $\Delta E \geq 2.3$. Therefore, we use 2.3 as a threshold to match the color bands to the reference symbols.

6.5 Final System

As mentioned before, this paper made an attempt to replicate the system mentioned in [Hu et al., 2015]. Due to time constraints, the system is unfortunately not finished. Like any attempt to create a system, some problems arise during the implementation phase of the system. These problems together with how they were solved are explained below.

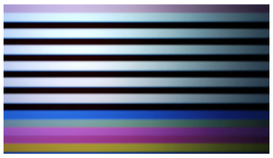

Figure 6: Single Frame containing the Calibration Packet

Transmitter

The transmitter was designed to send packets of three bits, these packets are encoding using the Reed-Solomon encoding defined in 6.2. Furthermore, the calibration packets are send after each 50 packets, to ensure correct synchronization between the receiver and transmitter. 6 shows such a calibration packet, here it is clear that the preamble is defined as white light followed by no light six times after each other followed by eight color bands indicating the eight symbols used in the modulation scheme.

One problem, we faced is that the color of the white light the preamble changed to a more blue color if the calibration packets were send more frequently. A reason as why this could happen is that the inter-symbol interference became too high, this would cause a color to change from the originally intended color.

Another problem, we faced was that the underlying R, G and B LEDs under the cap gave too much influence in the color of the color band. Our system made use of a diffuser to evenly distribute the lights. Another solution could be to increase the distance between the transmitter and receiver, unfortunately our system was not designed to have a distance of more than a couple centimeters between the transmitter and receiver. The latter approach is more recommended if the system were able to support more distance between the transmitter and receiver, since it is costly to attach a diffuser to each LED used in this system and a diffuser slightly decreases the brightness of an LED. Keeping in mind that the main purpose of an LED is to illuminate, a diffuser would work against this.

The designed system also made use of a tripod to hold the receiver still, this is because a slight trembling of the hands increased the inter-symbol interference which in turn caused double stripes in the received image and color bands with a different color than intended.

As last, the data transmission between the transmitter and receiver was held in a dark room to reduce the ambient lightning since ambient light may increase the inter-symbol interference and different color band interpretations during the data transmission.

Receiver

The receiver is the most complex part of the system, it must be able to recognize calibration packets, synchronize symbol colors and decode the data stream. The videos of the transmitting LED are captured using the Redmi Note 10 phone and processed offline. During the capturing of the video some problems arise with the use of the Redmi Note 10. When using the camera of the smartphone, the exposure time of the photodiodes must be shorter than a single symbol period. The symbol period used in the system is 1ms, so the exposure time must be shorter than 1ms. When using the receiver with an exposure time greather than a single symbol period, the inter-symbol interference is increased significantly. The increase in inter-symbol interference causes color bands to be received differently by the camera, the color of the color bands between different frames change too much to decode it reliably. Furthermore, the inter-symbol interference causes some color bands to be merged together which causes color bands that have an averaged color of the originally two separate color bands which in turn may cause a mismatch to the intended symbol. The way our system achieved an exposure time less than 1ms, is by setting the shutter speed of the Redmi Note 10 to 1/4000 and by setting the Exposure Compensation value to -4. The longer the shutter is open, the more light gets into the photodiodes, increasing the inter-symbol interference. By also setting the Exposure Compensation (EV) value to -4, the shutter speed is further decreased to ensure a exposure time less than 1ms.

After the video is taken with the correct parameters, the captured image is processed offline. The frames are extracted from the video and the color bands are found and their respectively color is extracted. These colors are then matched with their symbols and the message is decoded. A problem the system seemed to be showing is that the colors of the color bands are not consistent, they differ significantly during multiple frames. This makes it hard to match the color bands to their symbols, which limits the goodput of the system. The receivers exposure time may have something to do with this. Even though the system ensures an exposure time under one symbol period, there appears to be too much inter-symbol interference. This interference may also be caused by the symbol frequency, when the symbol frequency is increased the width of a color band decreases. This makes it hard to decode the incoming color band. Due to time constraints, this study has not investigated this problem.

7 Conclusions and Future Work

In this paper, we have designed and analysed a Tri-LED-to-Rolling-Shutter-Camera Communication System. The system makes use of Color Shift Keying to increase the data rate of the transmitted data, which in turn is captured and demodulated (offline) using a Redmi Note 10 equipped with a Rolling Shutter Camera and a laptop. We discussed and addressed three challenges when using Tri-LED-to-Camera Communication where the receiving camera is a Rolling Shutter Camera:

- Human perceivable Color Flickering
- Inter-Frame Data Loss
- Receiving Rolling Shutter Camera Diversity and Color Space differences

We have found a hypothesis that a lower CSK-Modulation method, which makes use of a lower number of colors, can achieve a much higher goodput. Whereas a higher CSK-Modulation can achieve a much higher throughput since it makes use of more color patterns.

We have shown in this paper that a much higher data rate is achievable using Color-Shift Keying as a modulation method for data transmission rather than On Off Keying or Frequency

Shift Keying modulation schemes. However, our designed system will only work when the transmitter and receiving camera are in close proximity in a dark environment. This could be improved to feature Tri-LED-to-Camera Communication for farther distances between the transmitting LED and receiving camera and to be more resistant to ambient lightning. Furthermore, our system has not optimized the Constellation points for Rolling Shutter Cameras, by optimizing the CSK Constellation design/points, a much higher data rate could be obtained since this optimized CSK Constellation design will reduce the inter-symbol interference. As explained earlier, due to time constraints, the system has not been finalized. There is still too much inter-symbol interference in the system present, which makes the demodulation error too high. A future implementation of this system must make sure to reduce this inter-symbol interference before implementing other features.

A trend that modern mobile devices are setting nowadays, is the narrow-bezel or no-bezel smartphones. The traditional sensors (fingerprint sensors, light sensors, and cameras) which were deployed on the mobile phone are now placed under the screen. A recent study [Ye and Wang, 2021] has enabled through-screen visible light communication using a transparent OLED screen. Although the highest data rate in [Ye and Wang, 2021] is about 2.6 kb/s which is 3.4x higher than traditional CSK they have found that the main bottleneck is the low speed of the minimum sampling time of the color sensor they used. This data rate could be improved if a rolling shutter camera is used as a receiver.

References

[Chan et al., 2017] Chan, C.-L., Tsai, H.-M., and Lin, K. C.-J. (2017). Poli: Long-range visible light communications using polarized light intensity modulation. page 109–120.

[Danakis et al., 2012] Danakis, C., Afgani, M., Povey, G., Underwood, I., and Haas, H. (2012). Using a cmos camera sensor for visible light communication. In *2012 IEEE Globecom Workshops*, pages 1244–1248. IEEE.

[Faruque, 2017] Faruque, S. (2017). *Frequency Shift Keying (FSK)*, pages 57–67. Springer International Publishing, Cham.

[Hu et al., 2015] Hu, P., Pathak, P. H., Feng, X., Fu, H., and Mohapatra, P. (2015). Colorbars: Increasing data rate of led-to-camera communication using color shift keying. In *Proceedings of the 11th ACM Conference on Emerging Networking Experiments and Technologies*, CoNEXT '15, New York, NY, USA. Association for Computing Machinery.

[Julian Randall and Burri, 2007] Julian Randall, Oliver Amft, J. B. and Burri, M. (2007). Luxtrace: indoor positioning using building illumination.

[Kuo et al., 2014] Kuo, Y.-S., Pannuto, P., Hsiao, K.-J., and Dutta, P. (2014). Luxapose: Indoor positioning with mobile phones and visible light.

[Landis, 1954] Landis, C. (1954). Determinants of the critical flicker-fusion threshold. *Physiological Reviews*, 34(2):259–286.

[Ly et al., 2020] Ly, B., Dyer, E., Feig, J., Chien, A., and Bino, S. (2020). Research techniques made simple: Cutaneous colorimetry: A reliable technique for objective skin color measurement. *The Journal of investigative dermatology*, 140:3–12.e1.

[Michal, 2021] Michal (2021). Flickering led lights effect – what causes led lights to flicker?

[Monteiro and Hranilovic, 2012a] Monteiro, E. and Hranilovic, S. (2012a). Constellation design for color-shift keying using interior point methods. In *2012 IEEE Globecom Workshops*, pages 1224–1228. IEEE.

[Monteiro and Hranilovic, 2012b] Monteiro, E. and Hranilovic, S. (2012b). Constellation design for color-shift keying using interior point methods. In *2012 IEEE Globecom Workshops*, pages 1224–1228.

[Monteiro and Hranilovic, 2014] Monteiro, E. and Hranilovic, S. (2014). Design and implementation of color-shift keying for visible light communications. *Journal of Lightwave Technology*, 32(10):2053–2060.

[N.O. Tippenhauer and Mangold, 2012] N.O. Tippenhauer, D. G. and Mangold, S. (2012). Toys communicating with leds: Enabling toy cars interaction.

[Rajagopal et al., 2014] Rajagopal, N., Lazik, P., and Rowe, A. (2014). Visual light landmarks for mobile devices. In *Proceedings of the 13th International Symposium on Information Processing in Sensor Networks*, IPSN '14, page 249–260. IEEE Press.

[Roberts, 2013] Roberts, R. D. (2013). Undersampled frequency shift on-off keying (ufsook) for camera communications (camcom). In *2013 22nd Wireless and Optical Communication Conference*, pages 645–648. IEEE.

[Sharma, 2002] Sharma, G. (2002). *Digital Color Imaging Handbook*. CRC Press, Inc., USA.

[Tianxing Li and Zhou, 2015] Tianxing Li, Chuankai An, Z. T. A. T. C. and Zhou, X. (2015). Human sensing using visible light communication.

[Tsai et al., 2014] Tsai, H.-M., Lin, H.-M., and Lee, H.-Y. (2014). Demo: Rollinglight - universal camera communications for single led. page 317–320.

[Ye and Wang, 2021] Ye, H. and Wang, Q. (2021). Spiderweb: Enabling through-screen visible light communication. In *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*, SenSys '21, page 316–328, New York, NY, USA. Association for Computing Machinery.

[Yoon et al., 2019] Yoon, S., Lee, K., Cha, J., Mariappan, V., Young, K., Woo, D., and Kim, J. (2019). Ieee standard for local and metropolitan area networks—part 15.7: Short-range optical wireless communications. *IEEE Std*, pages 1–407.

[Yu-Lin Wei and Lin, 2017] Yu-Lin Wei, Chang-Jung Huang, H.-M. T. and Lin, K. C.-J. (2017). Celli: Indoor positioning using polarized sweeping light beams.

[Zhao Tian, 2020] Zhao Tian, Charles J. Carver, Q. S.-M. R. A. Q. L. X. Z. (2020). Polartag: Invisible data with light polarization.

[Zhice Yang, 2015] Zhice Yang, Zeyu Wang, J. Z. C. H. Q. Z. (2015). Wearables can afford: Light-weight indoor positioning with visible light.