

Visual Light Landmarks for Mobile Devices

Niranjini Rajagopal, Patrick Lazik, Anthony Rowe

Electrical and Computer Engineering

Carnegie Mellon University

Pittsburgh, USA

{niranjir, plazik, agr}@andrew.cmu.edu

Abstract—The omnipresence of indoor lighting makes it an ideal vehicle for pervasive communication with mobile devices. In this paper, we present a communication scheme that enables interior ambient LED lighting systems to send data to mobile devices using either cameras or light sensors. By exploiting rolling shutter camera sensors that are common on tablets, laptops and smartphones, it is possible to detect high-frequency changes in light intensity reflected off of surfaces and in direct line-of-sight of the camera. We present a demodulation approach that allows smartphones to accurately detect frequencies as high as 8kHz with 0.2kHz channel separation. In order to avoid humanly perceivable flicker in the lighting, our system operates at frequencies above 2kHz and compensates for the non-ideal frequency response of standard LED drivers by adjusting the light's duty-cycle. By modulating the PWM signal commonly used to drive LED lighting systems, we are able to encode data that can be used as localization landmarks. We show through experiments how a binary frequency shift keying modulation scheme can be used to transmit data at 1.25 bytes per second (fast enough to send an ID code) from up to 29 unique light sources simultaneously in a single collision domain. We also show how tags can demodulate the same signals using a light sensor instead of a camera for low-power applications.

Keywords—Visual light communication, Indoor localization, Wireless communication

I. INTRODUCTION

Inside buildings, light bulbs are pervasive, have ample access to power, and are often ideally positioned for sensing applications. In this paper, we introduce a technique for transmitting data from solid-state luminaries, used for interior ambient lighting, to mobile devices like smartphones in a manner that is imperceptible to occupants. One compelling application of this approach is the ability to transform lighting sources into landmark beacons that can be used for indoor localization. If each luminaire could uniquely identify itself to nearby mobile devices, it would be possible to easily distinguish between rooms and areas within a room that are illuminated by different lights. Since visual light is completely blocked by solid obstructions like walls, this approach is well suited for semantic localization (position, rather than geographic), which remains a challenge for many existing ranging and localization systems.

In general, Visual Light Communication (VLC) holds the promise of providing an extremely high data-rate and low-cost network link. VLC systems have the potential for high signal-to-noise ratios (greater than 50dB), are not regulated by the FCC, and can be contained easily within walls providing a high degree of spatial diversity. However, the focus of this paper is not on creating a high-speed visual link, but rather

developing a visible light channel that can operate with off-the-shelf mobile devices. The proposed approach is a software solution for current mobile devices with cameras, and does not require any additional receiver hardware.

Beyond location landmarks, we see our approach being applicable to “invisible QR codes”, embedding digital tags into visual multimedia streams and physical-location-based security systems. We also see this technology being applicable to existing electronic devices that might already have LEDs that can be programmed to transmit data to mobile phones without requiring additional hardware.

As part of the Energy Independence and Security Act of 2007 and multiple Environmental Protection Agency initiatives, incandescent light bulbs are slowly being phased out across the United States and internationally for efficiency reasons. In residential and commercial sectors, lighting alone accounts for more than 17% of the total energy consumption [1]. Modern LED lighting drivers already provide a DC power supply as well as digital electronics responsible for maintaining color and brightness consistency, along with dimming capabilities. Therefore, the additional cost of adding our proposed modulation scheme and even wireless communication capabilities would be minimal. It stands to reason that there is an opportunity to use networked control and added functionality within lighting retrofits to further incentivize adoption and improve efficiency.

Utilizing existing lighting for communication is challenging for three main reasons: First, the system should be flicker-free. Second, since most mobile devices no longer have high-speed optical sensors (like IR receivers), we would like to decode data using available cameras that typically only operate at tens of hertz. Finally, due to the dense configurations of indoor lighting, we require the ability to support scalable multiple-access among transmitters. The main contribution of this paper is a practical Frequency Shift Keying (FSK) modulation approach that can be decoded using existing smartphones in a scalable manner by supporting multiple access.

LED light output, efficiency and lifetime are predominantly a function of the driving current and temperature of the silicon substrate. In order to maximize bulb lifetime and be able to compensate for output loss due to aging, most manufacturers limit the current flowing into the bulb, usually by means of pulse width modulation (PWM). Since turning an LED completely off by On-Off Keying (OOK) at camera capture frame rates would be easily detectable as a flicker, our approach uses FSK of a high frequency PWM signal to transmit bits.

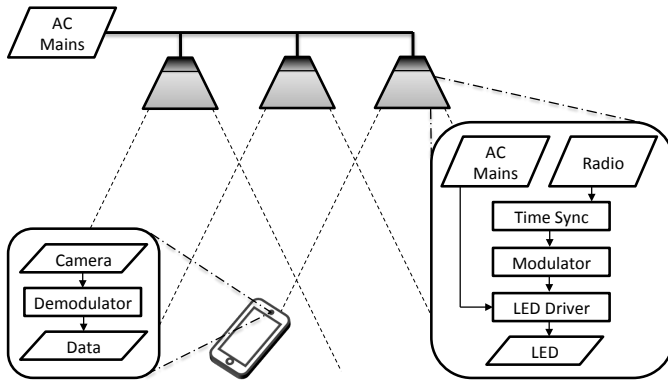


Fig. 1: System overview

In order to detect high-frequency changes in lighting, we exploit the fact that most modern CMOS cameras have a rolling shutter that does not expose the entire image simultaneously. Instead, the sensor pipelines the data transfer with the pixel exposure. As explained in more detail in Section III-A, this effectively means that a light pulsing at a frequency much higher than the frame capture time will illuminate only some rows at a time, producing bands in the image. We can detect the frequency of these bands in the image frequency domain, and infer the frequency of the pulsing light. We use this relationship to construct an FSK demodulator. Devices that are able to directly sample high-frequency PWM lighting signals, like photodiode detectors, can also decode this same signal. This means that the same system can communicate with low-power tags equipped with just a light sensor. Unfortunately, smartphone ambient light sensors are optimized for dynamic range and typically do not have the required frequency response to decode high speed data.

To support dense areas of unique transmitters, we use different frequencies to aid in multiple access. The receiver can process the incoming video data and decode all frequencies simultaneously. When available, time synchronization allows every node in an area to simultaneously transmit a common preamble, which makes synchronization with the start bit of the VLC data packet significantly more robust and leads to improved overall bit error rates (BER).

One of the major challenges is operating in a noisy channel that is polluted by background image data as well as ambient light. Camera parameters such as exposure and focus play a significant role in improving the ability to detect a signal. Through experimentation on a prototype system, we design and evaluate various modulation techniques and camera control algorithms for improving bit detection. The final modulator uses BFSK with 50 ms symbol lengths and is capable of transmitting 1.25 bytes per second reflected from illuminated surfaces and in direct LOS with reasonably high levels of background noise. Though quite slow as a data transport, this is fast enough to act as a beacon signalling system even when devices are carried by hand.

II. RELATED WORK

In this section we first examine existing research related to VLC, followed by work on the topic of indoor localization.

We then comment upon two systems which use VLC for localization.

A. Visual Light Communication

The vast majority of research on the topic of VLC focuses on employing it as a method for high-speed data transfer using specialized hardware. The favorable characteristics of using unlicensed spectrum at bandwidths of up to 100MHz [2], [3] with trichromatic (RGB) LEDs, or up to 20MHz with the more ubiquitous phosphorescent white LEDs [4], [5], make it an attractive contender for wireless communication. In 2006, the IEEE developed a draft VLC standard known as 802.15.7 [6]. In all of the previously referenced papers we find examples of using a variation of simple On-Off keying (OOK) for modulation. However, more complex schemes such as Quadrature Amplitude Modulation (QAM) and Discrete Multi-tone Modulation (DMT) / Orthogonal Frequency Division Multiplexing (OFDM) [7] are also discussed. Photodiodes are selected as detectors because of their large bandwidth. Due to the increased transmitter complexity of QAM, OFDM and DMT as well as the high bandwidth requirement, they are not well suited for camera communication systems. Simple OOK suffers from humanly perceivable flicker created by turning an LED completely off for bit sequences with several consecutive off periods. The major challenge in camera communication systems is reliably sending data without flicker at usable data rates.

Even though several iOS and Android smartphones are equipped with light sensors for automatically adjusting the back-light intensity of their screens, only Android allows direct software access to them. Furthermore, the sensors are usually CdS-based and optimized for dynamic range which does not provide enough bandwidth for most VLC applications. In [8] Danakis et. al. exploit the rolling shutter effect of a smartphone's CMOS camera to capture OOK modulated data from LEDs. The authors generate data sequences using Manchester encoding, resulting in multiple symbols being captured per frame. This allows for a transmission rate of up to 3.1kBd using 640x480 images at 20fps, but does not allow for simultaneous unique transmissions in the same collision domain and has questionable robustness when superimposed on noisy backgrounds. This solution requires close proximity with the LED and would likely not work for interior lighting. A further drawback is that the modulated signal can produce a human perceivable flicker from the transmitting LED. The authors alleviate this by imposing a DC bias on the signal, which in turn decreases its dynamic range and SNR at the receiver. This makes the scheme require significantly brighter lights and more complex driving hardware than our proposed approach.

VLC has been used to support mobile-to-mobile communication. COBRA [9] proposes a scheme for high-speed communication between small-sized screens and low-speed cameras with the transmit data refresh rate being half the frame rate. LightSync [10] resolves the problem of frame synchronization when the transmitting screen and receiving camera have variable frame rates. These approaches are for line-of-sight communication and actively filter out rolling shutter effects. VRCodes (NewsFlash) [11] takes advantage of the rolling shutter effect to decode visual tags displayed on LCDs.

The tags use multiple pixels of different colors, modulated at up to 120Hz to transmit data. The technology exploits the “flicker-fusion threshold” of the human eye to blend the tags into the background by rapidly flashing complimentary hues of color, still visible to a rolling shutter camera. This approach does not support multiple access beyond spatial diversity, making it less suitable for ambient lighting use-cases. This work does suggest how different color channels could be used to increase data throughput, while still keeping transmissions imperceptible to humans.

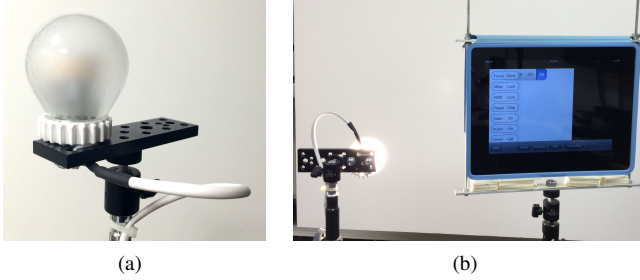


Fig. 2: (a) Cree 9.5W LED bulb (b) Preliminary setup

B. Indoor Localization with VLC

The related work on localization falls into two general categories, namely range-based approaches [12], [13], [14], [15], [16] and range-free approaches [17], [18]. Range-based approaches rely on measuring distances and/or angles with respect to known anchor points to compute a position based on propagation time of a signal. Due to the propagation speed of light, time-of-flight ranging is difficult with VLC, especially with cameras. Angle-of-arrival measurements are similarly difficult to attain reliably by using a single image sensor. Range-free localization approaches on the other hand typically attempt to match either synthetic or naturally occurring signatures to a particular location. For our system, we chose to employ this type of approach by having transmitters continuously send unique IDs over a VLC channel, which can be mapped to their known locations.

ByteLight [19] is a commercial effort that uses LED lights as visual landmarks. While we suspect that they exploit the rolling shutter effect of the camera and possibly modulate data, there is no detailed information about how the system operates and no available information describing their modulation technique, or the channel characteristics. At the time of publication, the lights were not commercially available, so we were unable to more closely inspect their design. In contrast, we present a modulation scheme and provide a detailed evaluation under various conditions. ByteLight also highlights that their underlying approach requires no inter-light communication. Our system can also operate without synchronization between the lights, but benefits if it is available.

CamCom [20] uses undersampled frequency shift OOK (UFSOOK) by encoding data at frequencies that are harmonics of the frame rate, and decoding data by processing the sub-sampled aliased frequencies. While this scheme works with both global as well as rolling shutter sensors, it operates at frequencies around 120Hz, which can cause stroboscopic flicker (see Section IV-E). For indoor navigation, the system

requires spatial diversity to support multiple access (data cannot be decoded from multiple transmitting lights mixed on a surface), which might not be suitable for certain lighting architectures. We believe it may be possible to mix aspects of UFSOOK with our approach, providing the best of both techniques.

III. SYSTEM ARCHITECTURE

Figure 1 shows the two main components of our system: a stationary VLC transmission infrastructure and multiple mobile receiving devices. Each transmitter uses a single, or multiple LEDs in an array to broadcast a unique ID. The ID is modulated onto the standard PWM signal used to drive commercial and residential LED lighting. As elaborated in Section III-B, in our implementation the modulated signal is generated by a low-cost micro-controller which drives a commercially available LED bulb through a simple MOSFET driver circuit. Due to the relatively low maximum signal frequency of 8kHz, our system is well within the attainable bandwidth of standard phosphorescent white LEDs [4] and standard MOSFETs. Each light transmits data using a Binary FSK (BFSK) modulation scheme and multiple transmitters are supported in a single collision domain by Frequency Division Multiple Access (FDMA).

The receiver is a common CMOS rolling shutter camera that captures the signal either through direct LOS, or reflected from surfaces. The data rate and received SNR depend on the camera’s capture speed and its resolution. Although we performed our evaluation using 720p at 30fps, lower settings such as 480p at 15fps can also be used. A more detailed discussion of how camera parameters impact communication can be found in Section IV. The captured signal is demodulated entirely in software running locally or off-board if raw video can be streamed to a server.

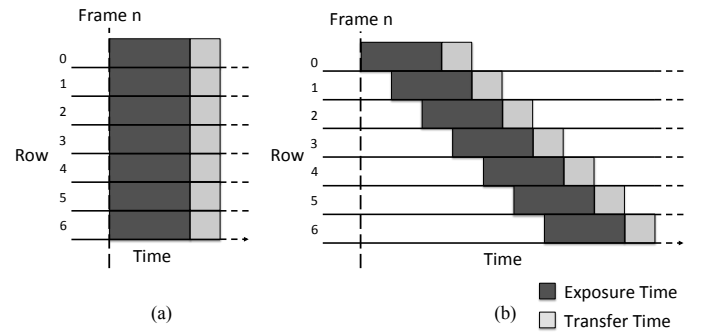


Fig. 3: Shutter mechanisms (a) Global (b) Rolling

A. Rolling Shutter

Figure 3 compares global and rolling shutter operation [21]. Global shutters, which are commonly implemented on CCD sensors (although CMOS variants exist), expose all pixels on the sensor simultaneously and gather incoming light over all pixels for the exposure time τ_e . After collection has stopped, the data is transferred. Rolling shutters on the other hand, consecutively expose and read-out individual rows of pixels in a pipelined fashion. The exposure is performed in rapid succession, producing adequate images for scenes with minimal motion. τ_e for a rolling shutter is defined as the

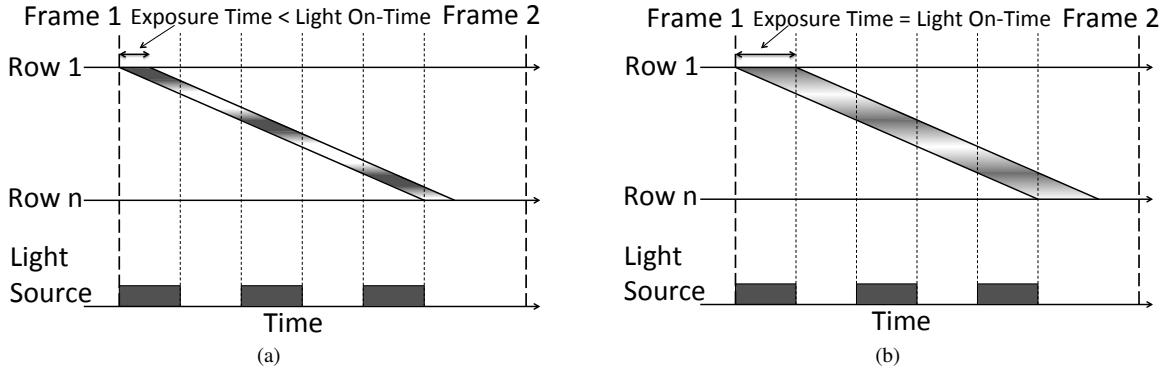


Fig. 4: Capturing a time varying light signal as a spatially varying image (a) Short and (b) Long exposure

time each individual row is exposed before being sampled. Though rapid motion or changes in lighting can result in significant geometric distortion of the captured images, rolling shutter technology is prevalent in CMOS sensors used in smartphones, tablets, and consumer computer peripherals due to the advantages of low power-consumption, complexity and cost.

We exploit the rolling shutter mechanism to capture a time-varying light as a spatially-varying image. As can be seen in Figure 4, an LED pulsed at a period τ_{LED} (less than the the frame duration) will produce bright and dark bands coinciding with rows exposed during the on-time τ_{on} and off-time τ_{off} of the LED respectively. The duty-cycle of the LED PWM signal determines the ratio of the height between the bright and dark bands. The number of bands per image is proportional to the LED's frequency $f_{LED} = 1/\tau_{LED}$, as well as the vertical resolution of the image. The height of each band is determined by f_{LED} and the row transfer time. The key phenomena is that the frequency of the image bands is proportional to the LED PWM frequency.

Figure 4 also shows that the smoothing between light and dark bands can be minimized by shortening τ_e . However, care needs to be taken to not under-expose an image which would decrease the SNR of the received signal. In order to maximize the dynamic range between light and dark bands, τ_e must be smaller than $\min(\tau_{on}, \tau_{off})$.

B. Prototype LED Luminaire

We based our prototype luminaire, shown in Figure 2(a), on a 9.5W Cree warm white (2700K) LED bulb [22], which outputs 800 lumens and is available for less than \$9. It has 80 white phosphorescent SMD LEDs, driven in series, that are arranged in a radial pattern inside the bulb. We replaced the bulb's power electronics with a simple MOSFET driver circuit, which is controlled from an external wireless micro-controller board based on the ATmega128rfa1 processor with integrated 802.15.4 radio. The ATmega allows us to program custom PWM patterns and also transmit data messages and synchronization pulses wirelessly. Our design includes a simple voltage divider circuit that allows the processor to precisely sample the 60Hz AC waveform of the mains power that can be used for time synchronization.

IV. VISUAL LIGHT COMMUNICATION

In this section, we describe the packet encoding scheme and the details of signal detection and demodulation.

A. Data Encoding

Each transmitter has a uniquely assigned frequency to represent *on*-bits, a shared preamble frequency and a shared *off*-bit frequency. Each packet, shown in Figure 7, contains one byte of data, along with 6 error correcting bits in a 14-8 Hamming format (two concatenated 7-4 codes), a preamble and a pilot symbol. The preamble is used to indicate the start of each data packet, while the pilot, which is identical to a transmitters *on* symbol, allows the receivers to measure the noise floor of each transmission.

Each transmitter in a single collision domain is allocated a unique frequency for transmitting its *on* symbol. In our implementation, the nodes are synchronized and use the same preamble frequency. However, if synchronization is not feasible, each node can have a different preamble frequency, at the cost of reduction in number of available frequency bands. The *off* symbol is identical among all transmitters and is broadcast at a frequency above the upper bound of the camera's frequency response, allowing efficient allocation of bandwidth and preventing the LEDs from turning *off* during the transmission of a 0-bit, hence eliminating flicker. In our tests, we use symbols that are at least as long as the frame capture time. Sub-frame symbols would support a higher data-rate, but require image background estimation. With a frame capture rate of 30fps and a frame duration of 33ms, the packet duration would be 528ms, giving a maximum location ID refresh rate of close to 2Hz. The periodic nature of the packet transmissions allows a receiver to synchronize to the infrastructure and save power between packet transmissions.

B. Frequency Detection

Unlike many communication systems, LED lights typically only allow binary PWM signal generation limiting a light to one frequency at a time. Additional hardware could make arbitrary waveform generation possible, but this would significantly increase cost and complexity.

1) *Detecting the LED frequency:* Figure 5(a) shows a zoomed in portion of a white surface illuminated by an LED pulsing at 2kHz. The contrast has been artificially increased

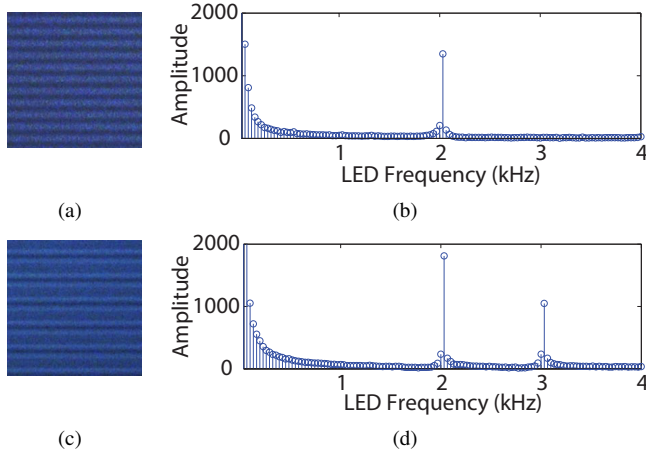


Fig. 5: (a) 2kHz image (b) FFT of 2kHz image (c) 2kHz and 3kHz from two different sources (d) FFT of mixed signal

for the sake of viewing. The 2-D Fourier Transform (FT) of an $M \times N$ image is computed as:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \exp^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})} \quad (1)$$

A key concept which simplifies the 2-D image analysis problem to a 1-D problem is that the rolling shutter will always result in a frequency purely in the vertical dimension of the image. The frequency of interest $F(0, \omega_1)$ does not change across the horizontal dimension. Hence our frequency analysis reduces from 2-D to 1-D in the following manner:

$$F(\omega_1) = \sum_{x=0}^{M-1} \left[\sum_{y=0}^{N-1} f(x, y) \exp^{-j2\pi(\frac{\omega_1 y}{N})} \right] \quad (2)$$

Figure 5(b) shows the 1-D FT of Figure 5(a) with a clear peak at 2kHz. Figure 5(d) shows the 1-D FT of Figure 5(c) which was captured by the simultaneous operation of one LED source at 2kHz and another at 3kHz.

2) Input-Output frequency translation: The first aspect of the system we studied is the relationship between LED frequency and the corresponding 1-D image frequency. Figure 6(a) shows the mapping between the input PWM frequency and the image frequency. The plots were captured using a 50% PWM duty-cycle although the duty-cycle has minimal impact on frequency detection. We see that the image frequency is a linear function of the LED frequency. In the bottom portion of the figure, we see the number of pixels in one period of the signal as a function of the LED input. Higher frequencies have fewer pixels, and hence higher attenuation.

We also analyzed images captured at resolutions 640x480, 1280x720, 1920x1080. Due to the nature of the rolling shutter, the resolution of the image only changes the ability of the system to capture very low-frequencies. The inter-line timing is constant at different resolutions, so the high-frequency response remains the same. The advantage of a higher resolution image is that it provides more samples for frequency estimation. The highest detectable input frequency is limited by the rolling shutter speed while the lowest frequency is limited by the camera resolution.

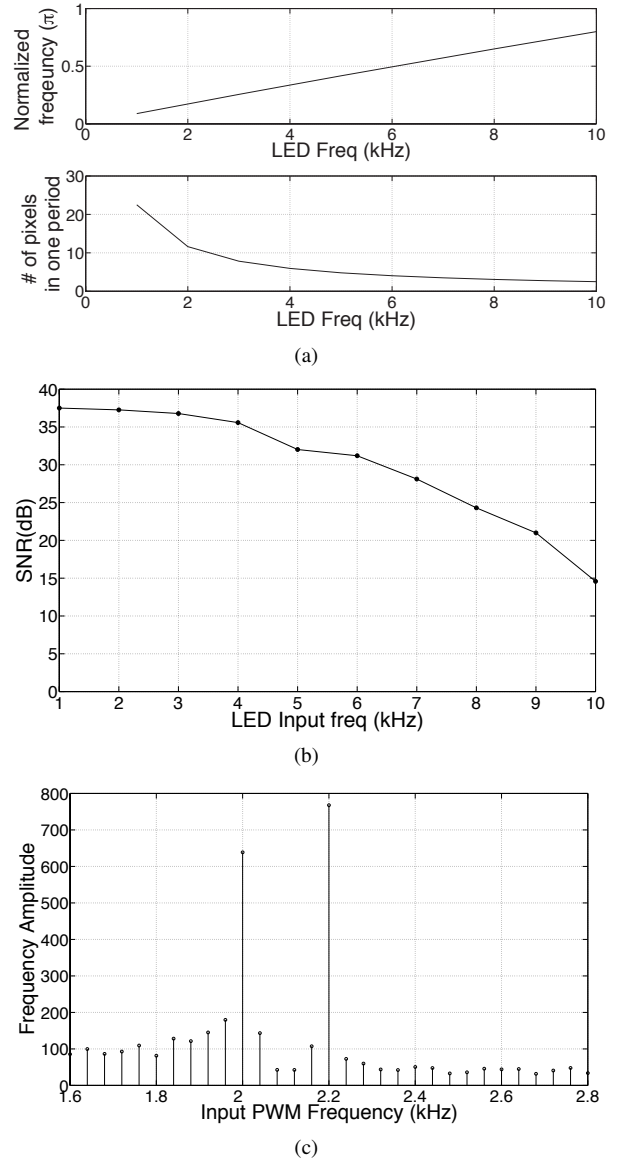


Fig. 6: (a) Mapping between LED frequency and image frequency (b) iPad 3 camera frequency response (c) Two PWM signals with 200Hz channel spacing

3) Frequency response of camera: Next, we look at the change in SNR over different LED PWM frequencies. We define signal-to-noise (SNR) ratio as the average power across a carrier frequency range when the carrier frequency is present as compared to the average power across the carrier frequency range when it is not present. Figure 6(b) shows the SNR in dB varying with the LED input frequency. We clearly see that the higher frequencies attenuate, but as we describe in the later sections, there is still typically enough illumination provided by overhead lights to operate as high as 8kHz or 9 kHz in practice. The lower frequency of operation is practically determined by the highest frequency perceivable by humans (see Section IV-E).

To prevent inter-symbol-interference (ISI), we need to ensure adequate spacing between symbols. Figure 6(c) shows a region of two signals that are spaced 200Hz apart. Since

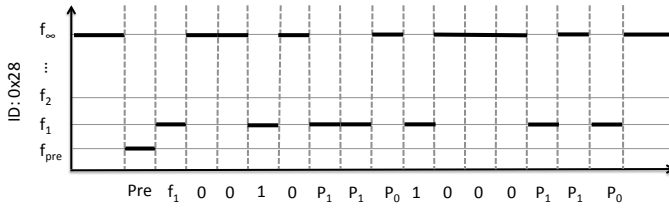


Fig. 7: Packet structure

the signal can occupy as much as 100 Hz of bandwidth, we conservatively set the channel spacing to 200Hz to reliably distinguish a signal from its neighbor.

C. Demodulation

Assuming a hypothetical receiver which is perfectly synchronized with the transmission infrastructure, and a symbol duration of one frame, each captured frame would contain exactly one symbol per received packet. However, since coordination between camera frames and the infrastructure is difficult, captured symbols usually spread across multiple frames.

1) *Stitching images - Sliding window approach*: We propose a sliding window approach where we first stitch all of the captured frames together into a single, long image as shown in Figure 8. We then window the data using a rectangular window of size equal to a single frame. By posing this restriction on the window length, we ensure that the frequency content of the background image is constant across all the windows. Each window of the image has three discontinuities: the top of the window, the bottom, and one in the middle caused by the concatenation of successive frames. To smooth the discontinuities, we apply two Hanning windows across each continuous segment.

A 1-D FT (as described in Section IV-B1) is then performed on the windowed image, reducing the image data to a one dimensional signal. Next, we compute the spectral power at the frequency of the preamble over a bandwidth of 100Hz. By sliding this window on a pixel-by-pixel basis, we obtain the change in spectral power over time. We can also slide the window in larger steps to reduce computation time.

2) *Preamble detection and decoding bits*: We detect the preamble by sliding the window as described above, and locating the window position corresponding to the highest power in the preamble frequency. Since the inter-symbol-spacing is known, the demodulator can now move the window downwards in steps equal to the height of a symbol length and precisely hit each data symbol of the captured packets. We use a binary threshold detector to distinguish the absence or presence of a symbol. The power corresponding to an *off* bit (or the absence of an *on* bit) is measured by the spectral power in the *on* frequency at the location of the preamble. The power corresponding to the presence of an *on* bit is measured by the spectral power in the *on* frequency, at the location of the pilot symbol. A threshold is then calculated to be the mean of these two values. The demodulator decodes the data symbols by comparing the power level at each possible symbol's frequency to the threshold. Once the demodulation is complete, the decoded bit sequences are checked for errors using a 14-8 Hamming decoder. Figure 9 shows a visualization

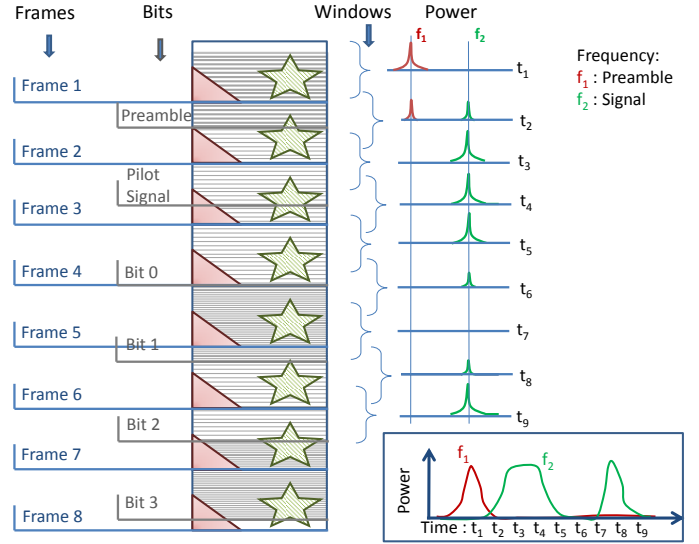


Fig. 8: Demodulating across frames

of the change in frequencies over the sequence of images during the demodulation process.

Rolling shutters cameras on mobile devices provide the ability to detect high frequency lighting signals, but also come with drawbacks:

- Data intensive: Large amounts of video data need to be captured and processed in real-time.
- The camera may capture substantial noise in the background of the VLC signal, which may be moving from frame to frame.
- Camera frequency response is limited. SNR degrades rapidly above 8kHz on current cameras.

We developed an iOS app for capturing and processing VLC data, based on Apple's AVCam sample application [23]. The app is capable of capturing up to 30 successive, uncompressed 720p frames at 30fps and processing them from memory. Due to the substantial amount of data being captured, we extract only a single color channel. The vast majority of digital image sensors use a Bayer color filter array, which includes twice as many green filters as red and blue ones in order to mimic the human eye's increased sensitivity to that wavelength [24]. Since our VLC signal content is highest in the green channel, we can disregard the other channels. Figure 10 shows a screenshot of our app.

D. Mobile Device Sensor Tuning

The automatic exposure control of modern CMOS cameras also significantly impacts the capture of the VLC signal. With automatic exposure set in continuous mode, the camera constantly adjusts its exposure based on the brightness of the scene, resulting in differences in the SNR over time. Although at the time of writing, iOS (version 7.02) does not allow the exposure to be set manually, it is possible to adjust it based on the luminosity of an arbitrary point in the view of the camera and locking it. Locking the exposure to the settings determined by the camera performs poorly, as the exposure is often too long, drowning out the VLC signal in ambient

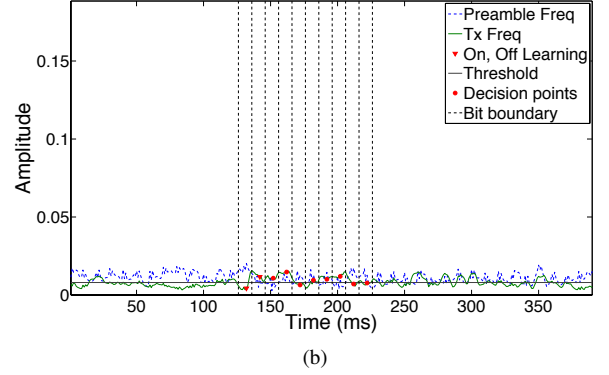
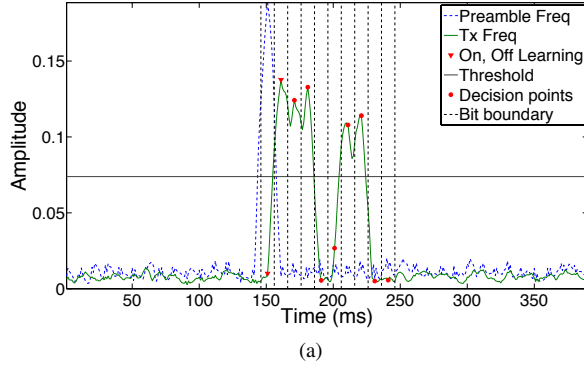


Fig. 9: (a) Visualization of demodulation (b) with high SNR input images (c) and low SNR



Fig. 10: iOS-based application

light. We developed an algorithm, which forces the camera to under-expose the image automatically, boosting the SNR of the VLC signal. First, we determine the brightest or dimmest areas (for decreasing or increasing exposure time respectively) in the current field of view and have the camera adjust its exposure accordingly. The app then continuously monitors the current exposure time and locks it once the target value is reached. The demodulator can also provide feedback on the induced change in SNR to this algorithm in order to find the optimal exposure. The improved performance due to the adaptive exposure algorithm is discussed in Section V-B. Android documentation includes API calls for manually adjusting the exposure but its functionality may not be implemented in all phone models.

A simple way of dealing with the background or subject captured in most images is to defocus the image. As shown in Figure 11, this has the effect of low-pass filtering, removing

all sharp edges present in the image, but retaining the bands created by the VLC signal. Similar to the limitations on exposure control, iOS does not allow manual setting of the focal range of the camera. A similar approach to our exposure control can be used by trying to automatically defocus the camera and then lock the focus. Again, Android already supports focusing APIs. A diffuser like scotch tape placed over the camera can also soften the image.

E. Flicker Reduction

There are two primary forms of flicker visible to humans: (1) *direct flicker* which are noticeable changes in light if viewing a static scene and (2) *stroboscopic flicker* which can be detected when objects in the scene are moving. In the lighting domain, researchers have rigorously investigated how occupants respond to lighting flicker and specify standards that avoid poor lighting that can cause physiological problems like headache and fatigue. In [25] and [26], the authors perform a user-study to define these limits and conclude that direct flicker can be avoided at frequencies above 60Hz, but that stroboscopic flicker can be irritating at frequencies below 2kHz. We use this as a parameter as the lower-bound of our frequency range.

One solution to reducing flicker, while keeping the minimum frequency low is to add a DC bias to the signal [8], keeping the LED continuously illuminated, and transmitting the VLC signal above the bias. The drawback is a significant reduction in dynamic range, resulting in a lower SNR at the receiver. For this reason, our carrier frequency is high enough to avoid flicker and our BFSK modulation scheme keeps the LED on even when 0 bits are being transmitted.

F. Data Rate

We achieve a data rate of 1.25 bytes per second for each transmitter. We believe this to be sufficient to transmit an ID for localization. However, the data rate can be improved in the following ways:

- 1) *Lower Symbol duration* : We use a symbol duration of 50ms. A lower symbol duration would enable a higher data rate but would be less robust to sampling jitter and have a lower SNR.

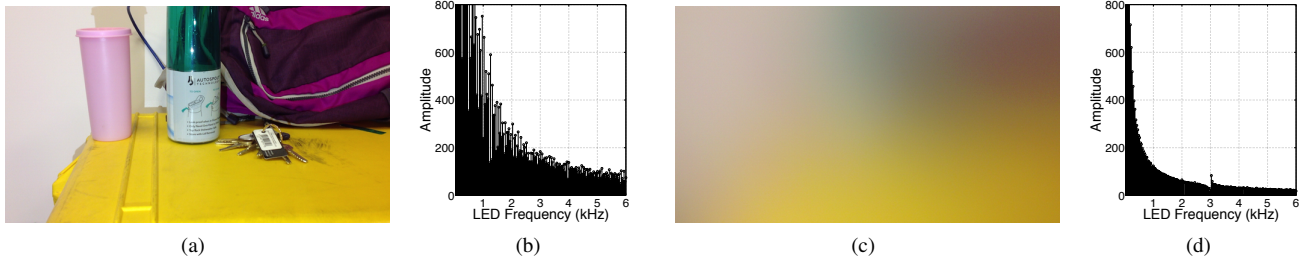


Fig. 11: Identical scene with 3kHz signal in sharp focus (a),(b) and defocused (c),(d)

- 2) *Increase frequency channels per transmitter* : We can obtain a higher data rate with a multiple FSK scheme, at the cost of supporting fewer transmitters in the same zone.
- 3) *Higher frame rate* : Newer phone cameras may support higher frame rate (iPhone 5s supports 120Hz burst capture). Our system can take advantage of this for a proportional increase in data rate.
- 4) *Arbitrary waveform input* : We designed our system with the constraint of a 2-state input to the LED. Higher data rate can be achieved if this constraint is relaxed, at the cost of a complex driving hardware which supports analog modulation schemes.

G. Tags

One benefit of our proposed approach is that low-power tags can also demodulate the signal. These can be aggressively duty-cycled and operate on micro-controller hardware. As a proof-of-concept, we attached a photo-transistor to the ADC of an ATmega128rfa1 micro-controller board. A photodiode is an ideal capture device as opposed to a CMOS camera, as it boasts a higher dynamic range, lower power-consumption, higher bandwidth, lower cost and higher noise immunity. While it was possible to capture over one second worth of 8-bit data at 15kHz using our sensor node, performing the demodulation locally is challenging, with only 16kB of RAM and 16MHz of processing power at our disposal. We therefore envision the tag compressing and uploading the captured data to a server back-end for demodulation and location lookup.

V. EXPERIMENTAL EVALUATION

We next evaluate the sensitivity of various design parameters in our modulation scheme in terms of how they impact real-world performance. We developed a hardware-in-the-loop experimental setup shown in Figure 2(b) which consists of a white backdrop, a set of our programmable LED luminaires with wireless communication and an iPad 3 mounted to a stand in order to capture consistent data. The luminaire and iPad are mounted 1 meter from the backdrop surface. For many of our experiments, it was important to be able to control the light output intensity without changing the duty-cycle of the LED. We achieved this by using a variable mechanical aperture control mounted in front of the light. In order to automate the data capture process, we have a version of our frame capture software running on an iPad that can be remotely triggered over the network. A data collection computer that has an interface to each LED luminaire coordinates synchronous frame capture along with LED data transmissions. We use a Matlab program

to collect image data stored locally on the iPad and process it against the transmitted test values. Using this automated process, we can rapidly experiment with different modulation schemes averaged over various light intensity levels and with different background images. For all of the experiments in this section, we used the front-facing camera on an iPad 3 collecting 720p HD video at 30fps with an f/2.4 fixed aperture. Each data point on the BER graphs was generated by the modulation and demodulation of 30 randomly generated test vector bytes.

It is to be noted that the experimental setup while evaluating multiple access (Section V-C), PRR in a realistic setup (Section V-D) and effect of camera motion (Section V-E), is different and has been explained in the respective sections.

A. Camera Sensitivity

As mentioned in Section IV-B3, we define signal-to-noise (SNR) ratio as the average power in a 100Hz bandwidth across a carrier frequency range when the carrier frequency is present as compared to the average power when it is not present. In Figure 12(a), we map these SNR values to illumination levels in Lux. A Lumen (lm) is the total flow of light emitted continually from a source in all directions. One lumen is equal to a small wax candle. Lux is a measurement of the intensity of light that falls upon a surface and is defined as lm per m^2 . On a clear sunny day, the sun provides approximately 50,000 Lux on the surface of the earth (1000 Lux on overcast days). Each of our luminaires provide a maximum of 700 Lux from 1 meter away which approximates indoor office lighting conditions.

Figure 12(b) shows the drop in SNR with distance when the light is in LOS of the camera, with the exposure and focus adjusted as mentioned in Section IV-D. The camera can decode data if the light is within 10m when in LOS. While the corresponding SNR of the photodiode is orders of magnitude larger, the drop in SNR with distance of the camera is slower than the photodiode.

As described in Section IV-D, the camera's exposure settings has a significant impact on the system's performance. Figure 12(c) shows that a smaller exposure gives a lower BER at a fixed illumination. However, if the exposure is set too low then there could potentially be no light detected. For this reason, our algorithm attempts to set the exposure as low as possible while maintaining a minimum image brightness level.

B. Modulation Parameters

Now that we have established a mapping from light intensity level to SNR values, we evaluate the impact of

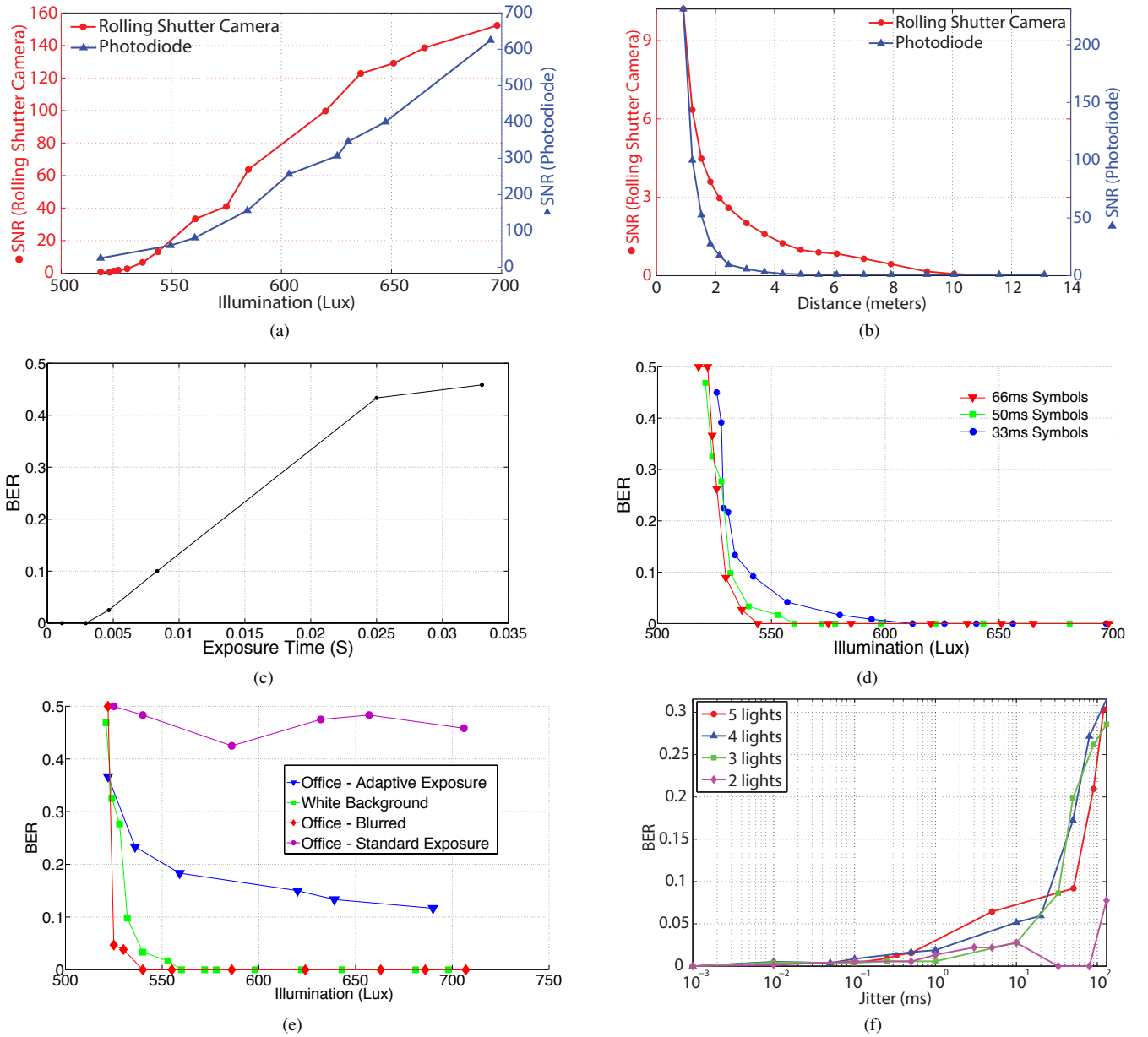


Fig. 12: (a) SNR vs transmit light intensity (b) SNR vs distance (c) Effect of exposure (d) Effect of symbol length (e) Effect of different backgrounds and tuning camera parameters (f) Jitter with multiple transmitters

transmission symbol length on BER. Since each data byte is encoded as 16 bits (as shown in Figure 7), we are limited to around 63 ms symbols for 1Hz updates. Figure 12(d) shows BER at increasing illumination levels with three different symbol lengths. Longer symbol lengths perform better due to larger windows of available data, and higher robustness to jitter in both the symbol sampling time, as well as, the camera frame capture time. The trade-off of using longer symbols is a lower data rate. We choose to use a 50 ms symbol length giving an effective data rate of 1.25 bytes per second. It is also worth noting that at lighting levels above 600 Lux, the BER rate is nearly zero.

One of the main challenges in the system is coping with background images since sharp edges in the image can generate frequency in similar ranges as our modulated data. We

first evaluate how a typically office environment scene impacts BER as compared to a uniform white background. *Office-Standard Exposure* and *White Background* in Figure 12(e) show the degradation in performance when operating on a non-uniform background. We next evaluate the effectiveness of controlling the camera parameters discussed in Section IV-D as they apply to realistic scenes. The *Office-Adaptive Exposure* line shows that by lowering the exposure while balancing the overall brightness of the image, the background artifacts normally found in the image are significantly reduced. The *Office-blurred* line shows how defocussing can drastically improve performance making the scene image almost as easy to decode as a uniform background image. Since the rolling shutter effect is completely independent of the image content, defocusing removes only the high-frequency artifacts that are

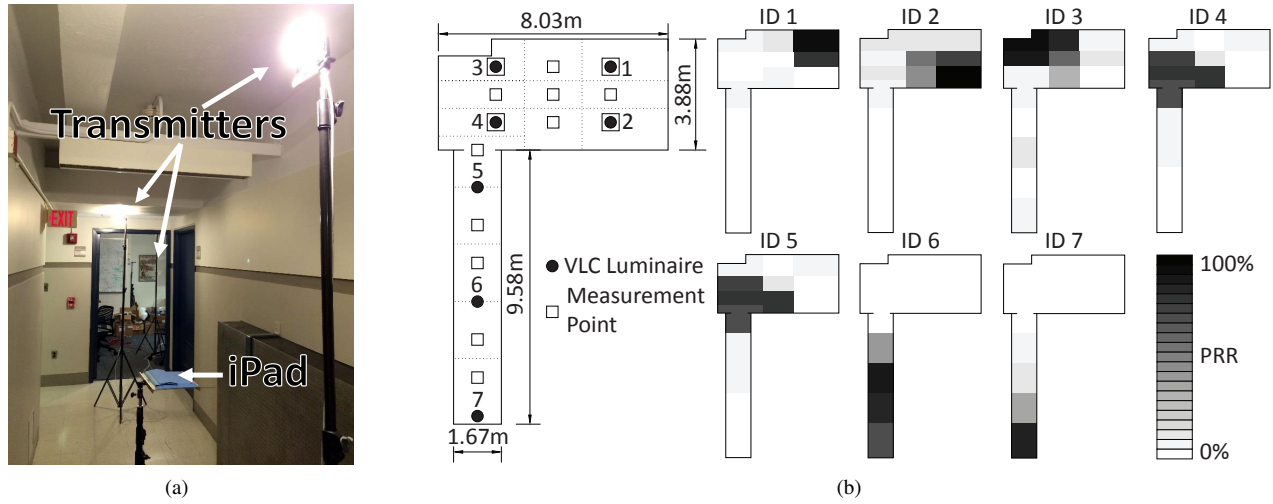


Fig. 13: PRR in a realistic environment (a) Experimental setup (b) Layout of room and hallway (left) and Results (right)

generated by the scene and leaves the horizontal banding complete intact.

C. Multiple Access

Next, we evaluate the ability of our system to scale with multiple simultaneous transmitters. With a frequency guard band of 200Hz and an operating range of 2kHz-8kHz, the system can support 29 different channels (plus 1 preamble channel for synchronization). This number decreases if synchronization is not available and the nodes transmit preambles on different channels. When operating with synchronized preambles, jitter in the preamble transmission impacts BER. Figure 12(f) evaluates the BER as we increase the jitter, under the worst-case scenario where all the interfering lights are co-located with the light under test. The errors remain almost unaffected for jitter less than a few milliseconds but increases for jitter in the order of symbol duration (50 ms). If the interfering lights are further away, their preamble would have a lower amplitude than the light under test and not affect the system performance. The synchronization of the preamble across the nodes is achieved using the 802.15.4 radios using a simple 1-hop flooding scheme. In realistic deployments, existing time synchronization protocols like Glossy [27] can be used. It is also worth noting that coloring schemes can be employed to allow frequency reuse among lights. Another benefit of shared preambles is that even in highly noisy environments (for example exposed to direct sunlight), the mobile device should still be able to detect the presence of a preamble even if the data is not receivable. This preamble can be used to narrow down the location of a mobile device.

D. Packet Reception Rate

To test the Packet Reception Rate (PRR) in a realistic setup, we installed seven luminaries 10cm below the ceiling, at various locations in our lab and adjoining hallway, as shown in Figure 13(a). The diagram on the left of Figure 13(b) shows the layout of the transmitters and measurement points. The points were chosen at key positions: directly below, surrounding and in between two luminaries. An iPad was mounted horizontally on a stand with the camera pointed upward, perpendicular

to the ceiling at a height of 1.20m. The heat maps on the right of Figure 13(b) show the PRR of individual IDs for 20 transmitted packets at each measurement point. In every case, the measurement point closest to a transmitter of a particular ID shows the highest PRR of $> 80\%$, with a consistent drop in PRR over distance. Generally PRR drops below 10% over a maximum distance of 6.5m from the transmitter if there are no obstacles blocking the signal. The heat map for ID 3 shows some packets being received over a longer distance, most likely due to false positive demodulation. If this setup were to be used as a positioning system, the incorporation of RSSI into determining the closest transmitter further dramatically increases the positioning accuracy, which in our tests jumped to 98%.

E. Effect of camera motion

In a realistic environment, the camera would be subject to motion due to the mobility of the user and shake in the user's hand. To test the effect of motion, we setup 4 lights in a corridor, each separated by 8 feet, with 9 test locations every 4 feet. The experiment was conducted by a user holding an iPad, moving across the corridor at three different speeds: stationary, walking (5kmph) and running (9kmph) while triggering the receiver when passing a test location. The PRR of each light was computed at each location by decoding 20 received packets. Figure 14 shows the result of the experiment with three main observations:

- 1) *A movement or swing in the camera does not deteriorate the performance if the user is stationary.* In section V-B we showed that defocussing an image is a desirable condition since it blurs out the subjects in the image, but does not affect the intensity of the VLC bands. A moving camera has a similar effect.
- 2) *The performance degrades when the user is running.* For each transmission, the energy level of the preamble and pilot symbols determine the threshold for decoding the data symbols. However, when the position drastically changes, the light intensity changes with distance causing the data symbols to have a different energy level as compared to the pilot. Decoding the data symbols with

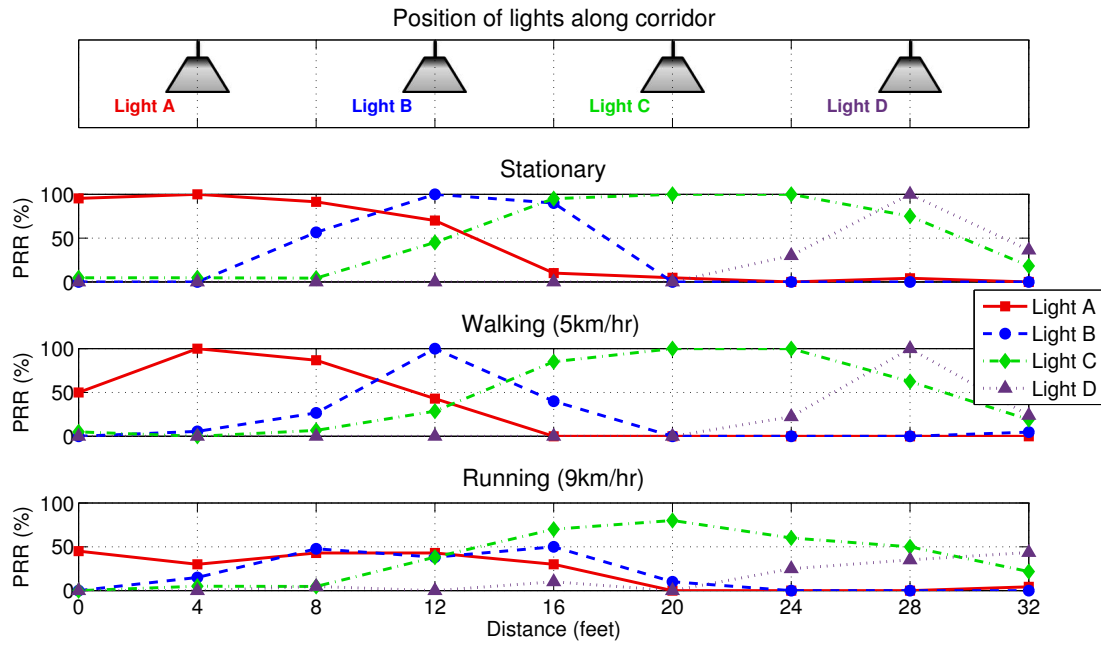


Fig. 14: Effect of pedestrian motion

the pilot energy level results in incorrect decoding of the bits.

- 3) *When the user's speed increases, the PRR spreads over distance.* The continuous data transmission causes a jitter in the time of preamble detection. While running, this jitter results in a large distribution of when the preamble is detected. With the lights transmitting their ID at a rate of 2Hz, and the user running at 9kmph, the user covers a distance of 4 feet during a single data transmission. With the lights separated by 8 feet in our experiment, this often resulted in detecting neighboring lights.

VI. CONCLUSIONS AND FUTURE WORK

In conclusion, this paper presented a technique for sending data from solid-state luminaries to rolling shutter cameras on mobile devices. One compelling use-case for this type of communication is the ability to use next-generation lighting sources as landmarks for indoor localization. We develop a prototype wireless LED lighting source and show that our approach is not only effective at communicating with smartphones, but can also be used to communicate with low-power embedded tag devices. Our prototype design is able to transmit at a data rate of 1.25 bytes per second on 29 channels concurrently to devices with CMOS cameras.

The focus of this work was predominantly on the communication channel between lights and mobile devices. This can immediately be used to localize devices by simply selecting the nearest landmark with the highest RSSI. However, for more accurate localization, one could develop tracking and estimation algorithms that can utilize multiple landmarks over time to automatically build maps and estimate more precise locations. Further innovation is also possible in terms of custom demodulation hardware on tag nodes that can significantly reduce their energy requirements. We believe that this approach has the potential to be an easy-to-deploy mechanism for retrofitting indoor spaces with intelligent lighting that can be

used to both conserve building energy and also help bootstrap future pervasive computing applications.

VII. ACKNOWLEDGEMENTS

This research was funded in part by the Intel Science and Technology Center on Embedded Computing, the Bosch Research and Technology Center in Pittsburgh and TerraSwarm, one of six centers of STARnet, a Semiconductor Research Corporation program sponsored by MARCO and DARPA. We would like to thank our reviewers and our shepherd Omprakash Gnawali for all of their great suggestions.

REFERENCES

- [1] U. E. I. Administration, "Annual energy outlook early release," 2013.
- [2] Y. Tanaka, T. Komine, S. Haruyama, and M. Nakagawa, "Indoor visible communication utilizing plural white leds as lighting," in *12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 2, 2001, pp. F-81–F-85 vol.2.
- [3] T. Komine and M. Nakagawa, "Fundamental analysis for visible-light communication system using led lights," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 1, pp. 100–107, 2004.
- [4] J. Grubor, S. C. J. Lee, K.-D. Langer, T. Koonen, and J. W. Walewski, "Wireless high-speed data transmission with phosphorescent white-light leds," in *33rd European Conference and Exhibition of Optical Communication - Post-Deadline Papers*, 2007, pp. 1–2.
- [5] J. Grubor, O. C. G. Jamett, J. W. Walewski, and K. d. Langer, "High-speed wireless indoor communication via visible light," in *ITG Fachbericht*, vol. 198, sept. 2007, pp. 203–208.
- [6] "Sg vlc project draft 5c," *IEEE Draft Std. IEEE P802. 15-08-0667-01-0v1c*, Sep. 2008.
- [7] H. Elgala, R. Mesleh, and H. Haas, "Indoor broadcasting via white leds and ofdm," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 3, pp. 1127–1134, 2009.

- [8] C. Danakis, M. Afgani, G. Povey, I. Underwood, and H. Haas, "Using a cmos camera sensor for visible light communication," in *IEEE Globecom Workshops*, 2012, pp. 1244–1248.
- [9] T. Hao, R. Zhou, and G. Xing, "Cobra: color barcode streaming for smartphone systems," in *Proceedings of the 10th international conference on Mobile systems, applications, and services*. ACM, 2012, pp. 85–98.
- [10] W. Hu, H. Gu, and Q. Pu, "Lightsync: unsynchronized visual communication over screen-camera links," in *Proceedings of the 19th annual international conference on Mobile computing & networking*. ACM, 2013, pp. 15–26.
- [11] G. Woo, A. Lippman, and R. Raskar, "Vrcodes: Unobtrusive and active visual codes for interaction by exploiting rolling shutter," in *IEEE International Symposium on Mixed and Augmented Reality*, ser. ISMAR '12, 2012, pp. 59–64.
- [12] P. Lazik and A. Rowe, "Indoor pseudo-ranging of mobile devices using ultrasonic chirps," in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, ser. SenSys '12. New York, NY, USA: ACM, 2012, pp. 99–112.
- [13] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan, "The cricket location-support system," in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (Mobicom '00)*. New York, NY, USA: ACM, 2000, pp. 32–43.
- [14] B. Parkinson and S. Gilbert, "Navstar: Global positioning system - ten years later," *Proceedings of the IEEE*, vol. 71, no. 10, pp. 1177 – 1186, oct. 1983.
- [15] T. Tanaka and S. Haruyama, "New position detection method using image sensor and visible light leds," in *Second International Conference on Machine Vision*, ser. ICMV '09, 2009, pp. 150–153.
- [16] A. Ward, A. Jones, and A. Hopper, "A new location technique for the active office," *IEEE Personal Communications*, vol. 4, no. 5, pp. 42 –47, oct 1997.
- [17] P. Bahl and V. Padmanabhan, "Radar: an in-building rf-based user location and tracking system," in *Proceedings of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '00)*, vol. 2, 2000, pp. 775 –784 vol.2.
- [18] K. Lorincz and M. Welsh, "Motetrack: a robust, decentralized approach to rf-based location tracking," in *Proceedings of the 1st International Conference on Location- and Context-Awareness (LoCA'05)*. Berlin, Heidelberg: Springer-Verlag, 2005, pp. 63–82.
- [19] A. Ganick and D. Ryan, "Self identifying modulated light source," no. Patent No. US20130026940, 01 2013.
- [20] R. D. Roberts, "Undersampled frequency shift on-off keying (ufsook) for camera communications (camcom)," ser. WOCC '13, 2013.
- [21] J. Nakamura, *Image Sensors and Signal Processing for Digital Still Cameras*, ser. Optical Science and Engineering. Taylor & Francis, 2005.
- [22] C. Inc., "Cree 9.5w (60w) warm white led bulb," 2009.
- [23] "Apple Inc. 2011. AVCam Sample Code. (March 2011). Retrieved January 29, 2013 from <https://developer.apple.com/library/ios/samplecode/AVCam/>."
- [24] B. Kisaécanin, S. Bhattacharyya, and S. Chai, *Embedded computer vision*, ser. Advances in Pattern Recognition Series. Springer-Verlag London Limited, 2009.
- [25] J. Bullough, K. S. Hickcox, T. Klein, and N. Narendran, "Effects of flicker characteristics from solid-state lighting on detection," in *Lighting Research and Technology*, 2011.
- [26] J. Bullough, K. S. Hickcox, T. Klein, and N. Narendran, "Detection and acceptability of stroboscopic effects from flicker," in *Lighting Research and Technology*, 2011.
- [27] F. Ferrari, M. Zimmerling, L. Thiele, and O. Saukh, "Efficient network flooding and time synchronization with glossy," in *Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on*, 2011, pp. 73–84.