Indoor Localization through Visible Light Characterization using Front-Facing Smartphone Camera

Charles Carver
Dept. of Physics & Eng. Physics
Fordham University
Bronx, NY
ccarver1@fordham.edu

Shela Wu
Courant Inst. of Math. Sci.
New York University
New York, NY
shela.wu@nyu.edu

Adriana Rogers
Dept. of Math. Sci.
Lewis & Clark College
Portland, OR
rogers.adriana@gmail.com

Matthew Stafford
Dept. of Comput. Sci. & Eng.
University of Buffalo
Buffalo, NY
mcstaffo@buffalo.edu

Dr. N. Sertac Artan
Dept. of Elect. & Comput. Eng.
New York Institute of Technology
New York, NY
nartan@nyit.edu

Dr. Ziqian Dong
Dept. of Elect. & Comput. Eng.
New York Institute of Technology
New York, NY
ziqian.dong@nyit.edu

***Abstract*—Research conducted in the field of localization with passive light, or using the intrinsic properties of light to determine a person's location, has seen increased growth in recent years. Specifically, fluorescent lights have been shown to exhibit distinct frequencies which can be recorded, along with their positions, for future lookup and positioning. Developments have been made in utilizing this phenomenon with a smartphone's high-resolution back-facing camera, however the constant flipping between the camera and the screen results in a poor user experience. In this paper, we propose an algorithm for extracting and analyzing both loop-shaped and tubular fluorescent lights. Similarly, we contribute an improved method for detecting frequency characteristics of unmodified fluorescent lights using a smartphone's front facing camera, therefore eliminating the need to constantly flip the phone.**

***Index Terms*—visible light localization, indoor positioning system, passive light source, localization with passive light, localization with visible light**

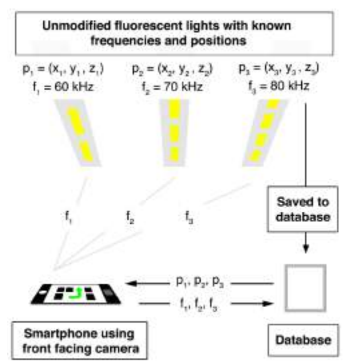


Fig. 1: Concept art showing how known fluorescent light positions and frequencies can be used for indoor localization without infrastructure modifications.

I. Introduction

Over the past few years, the prevalence of location based services, such as Global Positioning System (GPS), has risen in parallel with the increased adoption of smartphones. GPS, however, is generally unsuitable for indoor positioning as radio waves do not travel easily through solid objects, e.g., walls.

In an effort to overcome the limitations of GPS, visible light localization (VLL) systems have become prominent areas of research due to their high security and minimal interference qualities. These qualities can be attributed to the line of sight (LoS) characteristic that visible light exhibits, i.e., successful data transmission is only guaranteed if the receiver is within the field of vision of the transmitter. Unfortunately, most VLL systems require purposefully built light sources that are installed overhead which incur retrofitting costs. Fluorescent lights, however, have become an attractive candidate for VLL transmitters as they require no modifications to be used within a VLL system.

Due to how fluorescent lights operate, each light flickers at a unique characteristic frequency (CF) that is imperceptible to the human eye. One recent work introduces a VLL system named "LiTell" that utilizes unmodified fluorescent lights and a commercial smartphone for indoor positioning with an accuracy of 90.3%. [1].

The biggest drawback of LiTell is its reliance on a smartphone's back-facing camera. Given recent improvements in front-facing camera quality, we believe that there is potential for a front-facing camera, CF-based VLL system which would heighten user experience and create a more streamlined, easily adoptable VLL system.

This research, therefore, hypothesizes that, by using a smartphone's front-facing camera, we can sufficiently identify the characteristic frequency of a fluorescent light which may then be used for secure, indoor localization. We attempt this by examining a short video, taken by a smartphone's front-facing camera, to sample the frequency of any overhead lights. We then implement a video processing algorithm to identify the CF of both standard long-shaped and loop-shaped fluorescent lights by analyzing many frames within the video.

Section II begins by discussing existing VLL systems and comparing them. Required concepts are explained in Section III, including a background on how fluorescent lights and smartphone cameras operate. Our experimental setup and image processing algorithm are explained in Section IV. We analyze our results in Section V. We propose suggestions for future improvements in Section VI.

II. Related Works

A. Visible Light Localization

Utilizing visible light for indoor localization is a growing area of research as visible light provides a secure alternative to traditional RF-based systems. Many forms of VLL build upon existing visible light communication (VLC) schemes that allow for data transmission between LEDs and image sensors/photodiodes. These systems rely on modulating the transmitting LED, essentially turning it on and off at a rate that is imperceptible to the human eye. VLC systems with this basic design can also be used for positioning through means such as: proximity, fingerprinting, triangulation, and vision analysis [2].

In [3], Luxapose determined a smartphone's location with decimeter-level accuracy by modulating overhead LEDs and using an unmodified smartphone camera. Similarly, [4] focused on improving accuracy in a similar LED-to-smartphone localization setup.

B. Passive Source Systems

Recently proposed passive VLL systems are a promising alternative to active VLL systems [5]. While active VLL systems incur implementation costs, passive VLL systems do not as the light infrastructure remains unmodified. This is because passive VLL systems employ unaltered light sources, such as fluorescent lights or the sun, as location identifiers [5]. Notably, LiTell, an indoor VLL system, uses a mobile application to identify fluorescent lights with a smartphone's back-facing camera [1].

We investigate the usability of a smartphone's front-facing camera in a passive VLL system. Since the front-facing camera is capable of capturing JPEG images and videos, we test if frequencies can be extracted from either. Similarly, we assess various set-ups and smartphones to obtain the highest signal-to-noise ratio (SNR). In Section III-D, JPEG images are shown to be inoperative in passive VLL settings so we instead focus on using video for accurate CF detection.

III. Technical Background

This section explains the concepts integral to understanding localization using a smartphone and fluorescent lights.

A. Fluorescent Lights

Energy-efficient devices are generally favored, popularizing fluorescent lights within indoor office and retail settings. With a high level of light output accompanied by a low cost per lumen, fluorescent lights are an accessible source to be used within a VLL system [6].

Passive VLL systems can use the characteristic frequencies (>80 kHz) of fluorescent lights as unique transmitters. Within a pool of 500 FLs, less than 0.1% of lights had a difference in frequency less than 10 Hz [1], allowing for unique identification and viable transmitters for indoor localization.

B. Rolling Shutter Effect

Fig. 2: The rolling shutter effect captured by a Samsung Galaxy S8 over a modulated LED.

Most smartphone cameras use CMOS image sensors which sequentially expose sets of light sensors resulting in a banding effect as shown in Fig. 2 [7]. The rate at which one column of the image is exposed is called the sampling rate, calculated according to $W \times fps$, where W is the width of the image in pixels. The width of these bands vary depending on the modulation frequency which can be calculated through image processing techniques as explained in Section IV-B.

A FL's CF is much faster than the sampling rate of typical smartphone cameras, causing an aliasing effect. This effect occurs when a signal with frequency f is not sampled at $2f$, i.e., the Nyquist frequency, resulting in a false mirror image of the signal at frequencies below f [8]. To find the aliased frequency, we convert a time domain signal of the fluorescent light to the frequency domain by applying a fast Fourier transform (FFT). We then define our estimate of the true CF, f_e, as:

$$f_e = N f_s \pm f_a \tag{1}$$

where N is an integer greater than or equal to 0, f_s is the sampling frequency of the smartphone, and f_a is the aliased frequency [1].

C. Camera Settings

When taking short videos of the light, we alter the default camera settings (shutter speed, ISO, focus, and resolution) to optimally sample the lights and boost SNR. To detect fast frequencies, the shutter speed is minimized and the resolution is maximized. The ISO level was set to provide the maximum SNR, while not oversaturating the image. Finally, the video was blurred to mitigate the spatial artifacts caused by the lattice covers of fluorescent lights, as shown in Fig. 5.

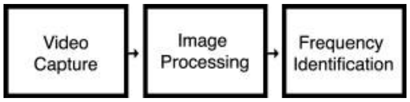


Fig. 3: A block diagram of our experiment.

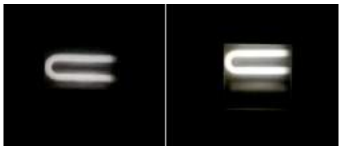


Fig. 4: Image quality of the Samsung S7, left, vs. the Samsung S8, right.

D. Compression Methods

RAW images were previously utilized by [1] for detail preservation but, to increase usability, we utilize the front-facing camera. Unfortunately, the front-facing camera only takes compressed JPEGs and videos. While JPEGs use a lossy compression algorithm that removes high intensity frequencies [9], MP4s are not compressed in the same fashion and are thus an alternative when using the front-facing camera.

IV. Experimental Setup

Our experimental setup is represented by Fig. 3 and is broken up into three steps: video capture, image processing, and frequency identification.

A. Video Capture

1) Camera Specifications: Initially, we tested three different phones: the Google Nexus 5, the Samsung Galaxy S7, and the Samsung Galaxy S8. The difference in image quality between the S7 and S8 can be visibly seen in Fig. 4. Certain camera settings are held constant throughout testing, including the shutter speed (1/14388.7s) and the video resolution (2880 x 2160).

2) Testing Set-up: We construct a controlled environment consisting of two differently shaped fluorescent lights: a pair of looped lights and a single straight light as shown in Fig. 5. Each light is turned on individually and all ambient light is excluded from the environment. Each light is also left on for at least 40 minutes to ensure frequency stability. A tripod is placed directly underneath the fluorescent lights so the smartphone lays roughly 131 cm from the ceiling. The smartphone is oriented in portrait mode so that the long edge of the fluorescent light fits within the short edge of the camera's field of vision.

Fig. 5: Two types of common light shapes, looped (left) and tube (right).

Once the camera is calibrated to the proper settings, data collection begins at ISO levels of approximately 800, 1200, 1600, and 2400. 90 second videos are recorded at each ISO level to ensure at least 60 seconds of usable data.

B. Image Processing

The image processing algorithm is mapped in Fig. 6.

A single video frame is flipped to portrait mode and converted to grayscale. The image is then thresholded to isolate the most intense region of light and morphologically opened and closed to remove any small pixel groupings. Boundaries are found across the image and the largest boundary, considered the main portion of light, is retained.

Rigid contours are drawn around the shape by identifying the points at which the boundary intersects bisectionional lines on both axes. When four indices are found, the FL is considered looped whereas two indices indicate a single rectangular FL.

The individual components are then extracted from the original image and converted back to grayscale. Each column of the component is looped over where the pixel intensity is summed and pushed to an array. The resulting intensity vector is fit to a 6th degree polynomial to remove lens vignetting [1]. Finally, the arrays from previous frames are concatenated.

C. Frequency Identification

To determine the aliased CF of the light, a fast Fourier transform (FFT) is applied to the vector of intensities. The number of samples for the signal is the size of the intensity vector.

The theoretical sampling frequency of the smartphone camera is calculated by multiplying the video's frame rate by the video's smaller dimension. To find the precise sampling frequency of the camera, however, we take a video of a modulated LED using the same camera settings as outlined in Section IV-A2. We use a function generator to power the LED at a range of frequencies and sample the resulting banding effect. We then apply image processing to determine the average width of each black and white band at each frequency tested. Finally, we run a fast Fourier transform over the vectors until the FFT results in a peak frequency at the original modulation frequency. For our experiment, the actual sampling frequency found over several trials was 64.93 ksps.

To preserve all subtle frequency components, we analyze the raw FFT output and look for any outlier frequencies that peak within the noisy signal. Windowing the Fourier transformation accentuates prominent peaks but occasionally diminishes unique frequency components that we are interested in observing.

V. Results

A. Image Processing

After analyzing our results, we find that there are consistent peaks across all three lights with various levels of intensity. As

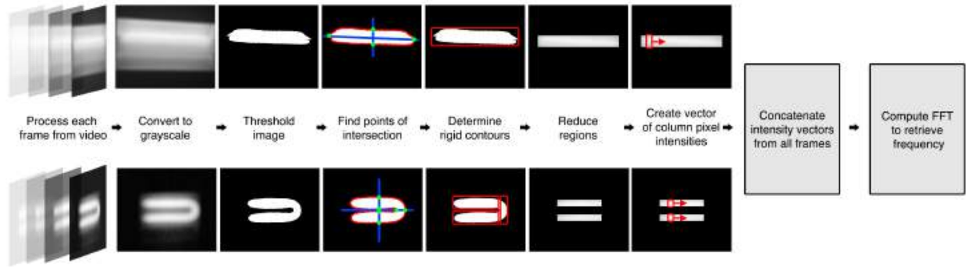
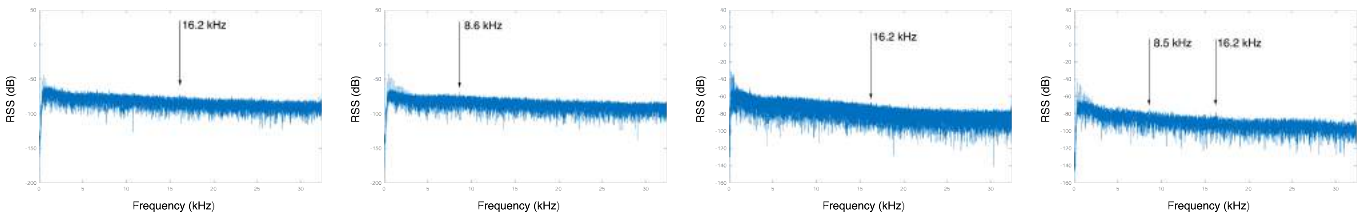


Fig. 6: Image processing algorithm for isolating regions of light.



(a) Light #1 (b) Light #2 (c) Light #3 (d) Light #3 after JPEG analysis

Fig. 7: Frequency distortions found across three different lights.

shown in Fig. 7., we can see that across the three lights, the RSS peaks at around 8.6 kHz and 16.2 kHz. The spectrum in Fig. 7d. was calculated from JPEG images and shows peaks at both 8.5 and 16.2 kHz in the same dataset. Since these frequencies are found within all three datasets, they are either erroneous side effects of our image processing algorithm/MP4 compression or simply frequencies common to all FLs.

Fig. 8. shows the dominant frequencies at ISO values of 801, 1196, and 2401 for one single light. A peak at 16.2 kHz is consistent among all three ISO levels, with the highest ISO level exaggerating consecutive peaks spaced roughly in increments of 4 kHz.

It appears that a higher ISO value skews the image processing and accentuates common frequencies while dampening unique frequencies. This is possibly due to how the image processing algorithm sums across the image's columns and adds up the pixel values. If the ISO value is lower, the small differences between dark and light bands are captured in the grayscale image and not simply washed out. Given our results, we find the optimal ISO value to be between 800 and 1200 with our given shutter speed and sampling rate.

B. Data Formats

We hypothesized in Section III-D that JPEG compression removes the high intensity frequencies that we require for this research. To evaluate this, we look at spectra differences between data formats captured from one light at an ISO value of 1201. We apply a Blackman filter to both sets of data and observe the resulting frequency spectrum. The JPEG data only showed a dominant frequency at 16.2 kHz whereas the video captured a unique frequency around 18 kHz. Consequently, we believe that high frequency data is, in fact, removed by JPEG compression but subtly retained in MP4 videos.

C. Frequency Uniqueness

Finally, we analyze the resulting frequency spectra obtained from our three experimental lights and emphasized the unique peaks in Fig. 9. We find that light #1 has two prominent peaks at 8.1 kHz and 22.1 kHz, light #2 has a prominent peak at 18.6 kHz (which is made more visible after windowing), and light #3 has a prominent peak at 2.4 kHz.

We believe that the SNR from light #3 was significantly higher due to the shape of the light. Light #1 and light #2 were both U-shaped lights and the image processing algorithm acted over the two parallel sections independently. Light #3, however, was one singular tube-shaped fluorescent light and consequently had a longer signal. Given Fig. 9, a longer signal in one frame results in a higher SNR.

Regardless of the differences in SNR, we have shown that there are unique frequency components between lights given no ambient light. We outline additional improvements in Section VI.

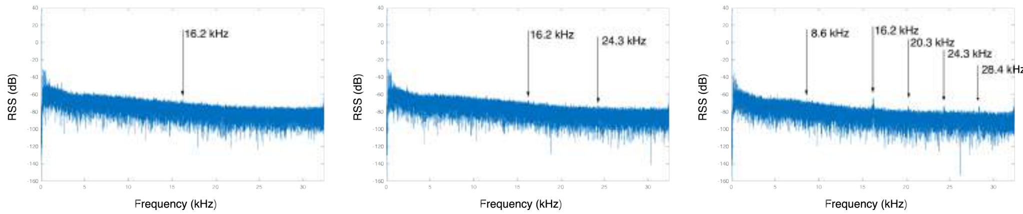
VI. Future Work

A. Smartphone

Typical smartphones shoot at 30 fps which is incomparable to the high characteristic frequencies of fluorescent lights. While videos provided a large amount of samples over a given period of time, the sampling rate is still too low. Finding a smartphone with a greater camera resolution could ensure a more suitable sampling rate.

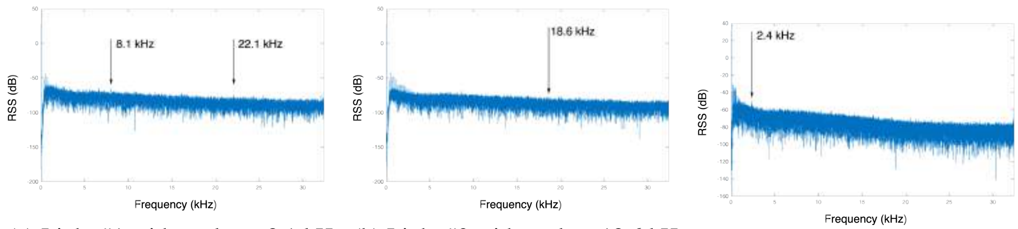
While testing Open Camera on various smartphones, we found that not all camera2 API features were consistently functional. Thus, we recommend developing an alternative application that allows for more accurate customization.

Videos and JPEGs from the front-facing camera are still fairly lossy. Other compression algorithms should be exam-



(a) ISO 801 (b) ISO 1196 (c) ISO 2401

Fig. 8: ISO's effects on frequency distortions for light #3.



(a) Light #1 with peaks at 8.1 kHz and 22.1 kHz (b) Light #2 with peak at 18.6 kHz (c) Light #3 with peak at 2.4 kHz

Fig. 9: Unique frequency peaks across three different lights.

ined, notably lossless compression techniques similar to RAW images that retain more of the original pixel data.

B. Experimental Set-Ups

In the future, having a measurement setup to record ground-truth CF values would allow for CF peak affirmation and aid in the development of further noise suppression techniques. Similarly, the effects of ambient light should be examined for real-world purposes.

VII. Conclusion

In this paper, we explore the feasibility of using a conventional smartphone's front facing camera to distinguish between unmodified fluorescent lights. We capture videos on a smartphone's front-facing camera and develop an image processing algorithm to extract aliased characteristic frequencies. Our evaluation shows that the front-facing camera has the ability to detect unique fluorescent light frequencies and offers significantly increased usability.

VIII. Acknowledgement

This research was supported by the National Science Foundation Research Experiences for Undergraduates program. We would like to thank all of the mentors, graduate student assistants, and research fellows who have provided their helpful insight and expertise that greatly assisted with our research. Similarly, we want to express our gratitude to the authors of [1] for helping us understand the complexities of their research.

This project is funded by National Science Foundation Grant No. CNS-1559652 and the New York Institute of Technology.

References

[1] C. Zhang and X. Zhang, "Litell: Robust indoor localization using unmodified light fixtures," in *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, pp. 230–242, ACM, 2016.

[2] T.-H. Do and M. Yoo, "An in-depth survey of visible light communication based positioning systems," *Sensors*, vol. 16, no. 5, p. 678, 2016.

[3] Y.-S. Kuo, P. Pannuto, K.-J. Hsiao, and P. Dutta, "Luxapose: Indoor positioning with mobile phones and visible light," in *Proceedings of the 20th annual international conference on Mobile computing and networking*, pp. 447–458, ACM, 2014.

[4] L. Li, P. Hu, C. Peng, G. Shen, and F. Zhao, "Epsilon: A visible light based positioning system.," in *NSDI*, pp. 331–343, 2014.

[5] Q. Wang and M. Zuniga, "Passive sensing and communication using visible light: Taxonomy, challenges and opportunities," *arXiv preprint arXiv:1704.01331*, 2017.

[6] "Lumens-per-watt comparison."

[7] C.-K. Liang, L.-W. Chang, and H. H. Chen, "Analysis and compensation of rolling shutter effect," *IEEE Transactions on Image Processing*, vol. 17, no. 8, pp. 1323–1330, 2008.

[8] M. Cerna and A. F. Harvey, "The fundamentals of fft-based signal analysis and measurement," *National Instruments, Junho*, p. 54, 2000.

[9] G. K. Wallace, "The jpeg still picture compression standard," *IEEE transactions on consumer electronics*, vol. 38, no. 1, pp. xviii–xxxiv, 1992.