INFRASTRUCTURE-LESS INDOOR LOCALIZATION USING LIGHT FINGERPRINTS

Shahab Hamidi-Rad, Kent Lyons, Naveen Goela

Technicolor Research†, Bay Area, USA

ABSTRACT

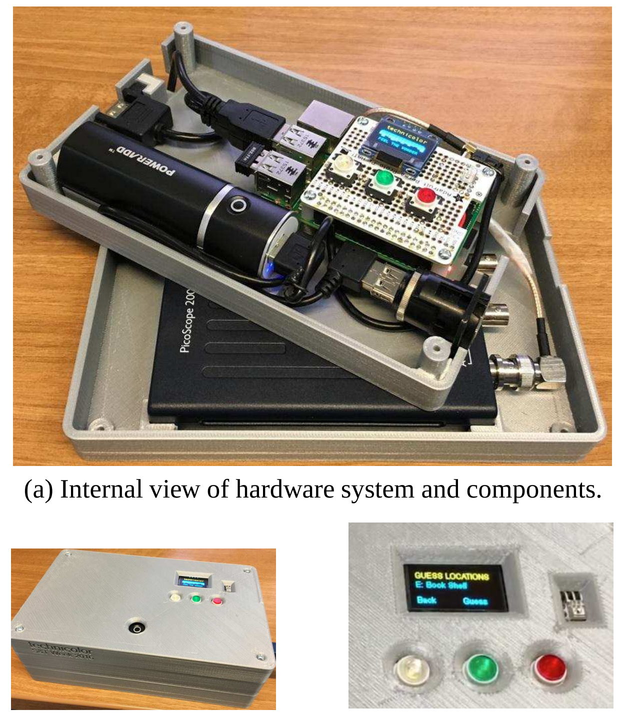
An infrastructure-less indoor localization system is proposed based on fingerprints of light signals acquired at high frequencies. In contrast to other systems that modulate lights, the proposed system distinguishes lights by learning from training samples. Due to slight differences in the electronic components used in the construction of compact fluorescent light (CFL) and light emitting diode (LED) bulbs, the optical signals emitted by each light bulb have slight differences with other light bulbs even within the same brand and model. Light signals are digitized with a fast and accurate analog-to-digital converter (ADC) at up to 1 mega-samples/second, segmented, and mapped into the frequency domain using the Fast Fourier Transform (FFT). Spectral features based on the FFT are filtered, normalized, and used as training data for supervised machine learning algorithms. Results are provided for two classifiers of varying complexity: (1) A k-Nearest Neighbor (KNN) classifier; (2) A Convolutional Neural Net (CNN) classifier. A hardware system for indoor localization was designed to analyze the performance of the classifiers. Under certain restrictions, results show that light bulbs may be identified with high accuracy without special infrastructure for modulation. Identifying a light bulb is meant to be synonymous with identifying its associated location.

***Index Terms*—** Light fingerprinting, sampling frequency, supervised learning, convolutional neural net, indoor localization.

1. INTRODUCTION

With the deployment of low-latency, high-speed 5G wireless communications, emerging technologies become possible due to connectivity within the "Internet of Things (*IoT*)" and "Internet of Signals". One important task to solve is indoor localization, detection, and classification of differentiated signals in an environment. This paper considers methods from machine learning to classify high-frequency light signals in indoor environments. Figure 1 depicts a light fingerprinting hardware system comprised of a high-frequency light sensor, analog-to-digital converter (ADC) scope, Raspberry-PI processor, and auxiliary components. Without additional infrastructure, the system collects data from different lights in a training phase, and distinguishes between multiple lights in a classification phase. The hardware components could be incorporated within a mobile device with the inclusion of light sensors with high-frequency sensing on the order of 1-MHz.

†Technicolor Research, 175 S. San Antonio Rd., Suite 200, Los Altos, CA, 94022, USA. Email Correspondence: {`shahab.hamidi-rad`; `kent.lyons`; `naveen.goela`}@`technicolor.com`.



(a) Internal view of hardware system and components.

(b) External view.

(c) Interface, light sensor.

Fig. 1. A hardware system for light fingerprinting and indoor localization is comprised of multiple components: Raspberry-PI processor, light sensor, ADC (scope), battery, user interface. The system with a light sensor is useful both for collecting training data, and for classifying lights. Additional infrastructure is not required.

Prior work regarding light fingerprinting relies on the explicit modulation of light transmitters in a visible light communication (VLC) framework [1–5]. Other accurate localization systems such as Epsilon or Luxapose utilize pulse-width modulation (PWM) or on-off keying to transmit identifiers, and use trilateration to localize devices [6, 7]. Without infrastructure, fingerprinting indoor locations was achieved via

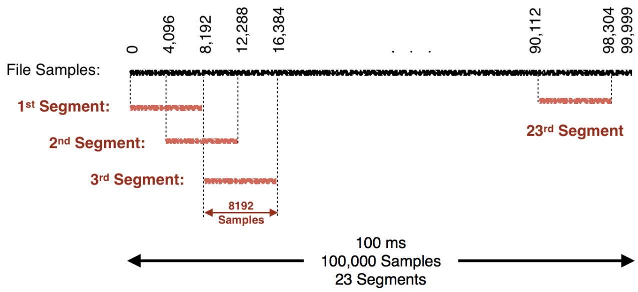


Fig. 2. Signal processing of light samples acquired at 100-ms intervals at a rate of 10^6 samples/second (1-MHz).

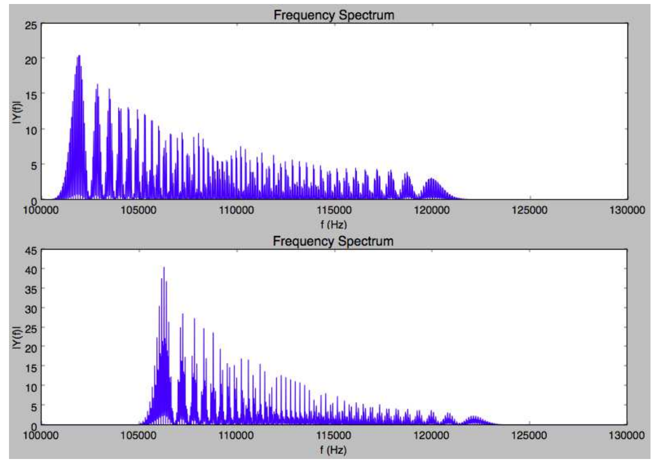


Fig. 3. *Top and Bottom*: The FFT magnitude response of two instantiations of a CFL light bulb (Ecosmart brand) in the [100-130] KHz range. The goal is to differentiate the switching pattern of the lights.

light intensity histograms in the Fiat-Lux system [8]. However, light intensity is not unique to individual lights, and is affected by several factors such as the distance between the sensor and light, room reflections, and variation of illumination over time of day. Another idea is to combine ambient sound, light, color, and motion via accelerometer to obtain a photo-acoustic signature of a location [9]. Recently, the IDyLL system was proposed for indoor localization using a fusion of signals from light sensors and inertial measurement units (IMUs) [10].

The approach presented in this paper is focused on fingerprinting individual lights based on their high-frequency fluctuations and switching patterns. The hypothesis explored via experiments is that each light provides a unique spectral signature if high-frequency sensing is available. High-frequency sensing capabilities beyond the KHz range may be present in future mobile phones. Concurrent to our research, authors of [11, 12] introduce an indoor localization system LiTell using unmodified light fixtures. In our approach, modern machine learning algorithms such as convolutional neural nets (CNNs) are harnessed to train multi-layer deep classifiers to distinguish light bulbs. Such classifiers do not rely on handcrafted signal features and have the potential to generalize to new inputs given enough training data.

2. SIGNAL ACQUISITION AND PROCESSING

2.1. Signal Acquisition

The acquisition of light signals depends on the characteristics of both the high-speed light sensor and the ADC. Our system incorporated a photodiode sensor (light-to-voltage TSL14S) with built-in pre-amplifier sensitive to light fluctuations and fast switching patterns. The analog signal of the sensor was connected to an 8-bits/sample ADC (PicoScope 2000) operating at $f_s = 10^6$ samples/second. The ADC was connected via a USB cable to a Raspberry-PI processer. Each light sensing phase of data collection lasted 100-ms in duration yielding $100,000$ samples in the time domain per data file.

2.2. Signal Processing– Time and Frequency Domain

Each data file containing $100,000$ samples per light bulb was further processed as displayed in Figure 2. The $100,000$ samples were divided into overlapping segments of length 8192 samples with 50% overlap, resulting in 23 consecutive overlapping segments per data file. An 8192-point FFT was computed per segment mapping the time domain signal to the frequency domain. The magnitude of the frequency response was retained (4096 distinct values). As shown in Figure 3, the magnitude of the samples in the frequency domain reveals differences in the high-speed switching patterns for two instantiations of CFL light bulbs (Ecosmart brand). Depending on the time of day, warm-up effects, and/or power-line effects, the frequency samples may be shifted or modified slightly. However, the aim is to learn the "shape" of the frequency response which is the invariant fingerprint.

2.3. Feature Extraction, Filtering, Normalization

Each data file containing 23 segments with corresponding 4096-length FFT magnitudes per segment was processed to extract feature vectors for experiments. In a first experiment, only 1400 magnitude values out of 4096 total were kept with associated frequency range $[20\text{-}190.9]$ KHz. In a second experiment, half of the 4096 magnitude values were kept with associated frequency range $[5\text{-}255]$ KHz. This yielded 2048 spectral features per segment. With our specific light sensor and 8-bits/sample ADC, incorporating higher frequency ranges did not improve classification accuracy significantly. The mean of the FFT magnitude values was subtracted, and all magnitude values were divided by the maximum value to normalize the feature vector for each light bulb. Normalization ensures that classifiers recognize the switching pattern of lights, and not extraneous acquisition biases.

3. SUPERVISED TRAINING OF CLASSIFIERS

The training and test data for CFL and LED light bulbs was collected at different times. Based on the training data and given labels for each light bulb, two main classifiers were trained of varying complexity[1].

3.1. k-Nearest Neighbor Classifier

The k-Nearest Neighbor algorithm is one of the simplest supervised machine learning algorithms, and its low-complexity implementation is ideal for the Raspberry-PI processor of our system, and mobile applications. A neighborhood parameter of $k = 3$ closest training examples was chosen based on standard cross-validation methods.

3.2. Convolutional Neural Net (CNN) Classifier

If a large amount of training data is available, a CNN classifier with multiple layers has the potential to generalize to new inputs and learn variations in the switching patterns of lights. In our controlled experiments, we were able to collect fingerprint data-sets of a size on the order of the well-known MNIST handwritten digits database. To show a proof-of-concept, our CNN classifier is comprised as follows: (1) Input convolutional layer; (2) Max-pool layer; (3) Fully-connected layer; (4) Output (soft-max regression) layer. The following parameter and input-output descriptions follow closely to those specified in the literature (e.g. TensorFlow open source models). The convolutional layer is inspired by state-of-the-art one-dimensional speech and two-dimensional image processing via deep learning.

For the first experiment, the input size to the CNN was 1400 spectral features. A convolutional kernel width of 251 was selected with kernel stride 1 and kernel depth 256 yielding a convolutional output volume of $(1400 \times 1 \times 256)$. A max-pool kernel width of 4 and stride 4 yielded a max-pool output volume of $(350 \times 1 \times 256)$. The fully-connected layer had input size $350 \times 1 \times 256 = 89600$ and output size 128. The final output layer had input size 128 and output of 6 classes. Before the final layer, a dropout probability of 0.3 was selected to avoid over-fitting, and a regularization factor of 0.001 was chosen. The batch size was 50 input vectors, and the CNN was trained for 100 epochs. The initial learning rate was 0.05 with a decay rate of 0.95. Initialization of CNN parameters was zero-mean with standard deviation 1.0.

Similarly, for a second experiment with input feature size 2048, the convolutional kernel width was 151 with stride 1 and kernel depth 128. The convolutional output volume was $2048 \times 1 \times 128$. After a max-pool layer with kernel width 4 and stride 4, the max-pool output volume was $512 \times 1 \times 128$. A fully-connected layer was chosen with input size $512 \times 1 \times 128 = 65536$ and output size 128. A final output layer had input size 128 and output size 8 representing 8 classes. Dropout was not utilized, but a regularization factor of 0.005 was selected. The batch size was 50 input vectors, and the CNN was trained for 100 epochs as in the first experiment. The initial learning rate was 0.05 with a decay rate of 0.95. Initialization of CNN parameters was zero-mean with standard deviation 1.0. All CNN classifiers were trained using the well-known back-propagation algorithm, minimizing a (cross-entropy) cost objective, and applying gradient descent.

[1]Alternate classifiers such as support vector machines (SVMs) yielded comparable accuracies for classification. However, given more training data, the multi-layer CNN architecture presented in this paper has a higher potential to learn complex data variations and generalize to new inputs.

4. EXPERIMENTS AND RESULTS

Several experiments with CFL and LED bulbs were conducted to analyze the performance of light fingerprinting for indoor localization. As described in previous sections, two main experiments are presented below for classifying *individual* light bulbs[2].

4.1. First Experiment: A Mixed Data-Set

In the first experiment, the data-set is comprised of 6 light bulbs including: 1 LED (Cree brand), 3 CFLs (Ecosmart brand), and 2 CFLs (G. E. brand). A total of 16 data files for training were recorded for each light bulb, yielding $16 \times 23 = 368$ training feature vectors per bulb. A total of 8 data files per bulb were recorded as a test-set after 2-days of separation between train and test samples. A total of $8 \times 23 = 184$ feature vectors per bulb were used for testing. Table 1 shows the confusion matrix results of the CNN classifier. A total accuracy of 97.10% is obtained. As a comparison, the k-nearest neighbor classifier achieved 93.11% accuracy (confusion matrix not shown).

	C1	E1	E2	E3	G1	G2
LED-C1	184	0	0	0	0	0
CFL-E1	0	184	0	0	0	0
CFL-E2	0	0	184	0	0	0
CFL-E3	0	0	0	184	0	0
CFL-G1	0	0	0	0	183	1
CFL-G2	0	0	0	0	31	153

Table 1. Confusion matrix for CNN classifier for 6 lights including 1 LED (Cree brand), 3 CFLs (Ecosmart brand), and 2 CFLs (G. E. brand). The train and test samples are taken with 2 days of separation. Total accuracy of 97.10%.

4.2. Second Experiment: Lights Of The Same Brand

In the second experiment, the data-set is comprised of 8 light bulbs which are all LEDs (Cree brand). A total of 40 data

[2]It is hypothesized that different *arrays of lights* (e.g., LED arrays) could also be distinguished via their combined group fingerprint.

files for training were recorded for each light bulb, yielding $40 \times 23 = 920$ training feature vectors per bulb. A total of 20 data files per bulb were recorded as a test-set on the same day as the training captures. In addition, a total of 20 data files per bulb were recorded as a test-set after 2-days of separation between train and test samples. A total of $20 \times 23 = 460$ feature vectors per bulb were used for testing in both cases. Table 2 shows the confusion matrix results of the CNN classifier for the test-set taken on the same day as the training captures. A total accuracy of 94.48% is obtained. As a comparison, the k-nearest neighbor classifier achieved 93.04% accuracy (confusion matrix not shown). Similarly, Table 3 shows the confusion matrix results of the CNN classifier for the test-set taken after 2-days of separation between training and testing. A total accuracy of 76.11% is obtained. As a comparison, the k-nearest neighbor classifier achieved 76.41%, an almost identical result in accuracy (confusion matrix not shown). The confusion matrix shows that 2 of the 8 LEDs caused most of the classification error.

	C1	C2	C3	C4	C5	C6	C7	C8
LED-C1	452	8	0	0	0	0	0	0
LED-C2	12	439	9	0	0	0	0	0
LED-C3	0	2	458	0	0	0	0	0
LED-C4	0	0	1	459	0	0	0	0
LED-C5	0	0	0	0	460	0	0	0
LED-C6	0	0	0	0	0	458	0	2
LED-C7	0	0	0	0	0	0	327	133
LED-C8	0	0	0	0	0	0	36	424

Table 2. Confusion matrix for CNN classifier for 8 instantiations of an LED light (Cree brand). The train and test samples were taken from the same day of data collection. Total accuracy of 94.48%.

	C1	C2	C3	C4	C5	C6	C7	C8
LED-C1	453	7	0	0	0	0	0	0
LED-C2	23	432	5	0	0	0	0	0
LED-C3	0	0	155	0	305	0	0	0
LED-C4	0	0	16	444	0	0	0	0
LED-C5	0	0	0	0	460	0	0	0
LED-C6	0	0	0	0	23	436	0	1
LED-C7	0	0	0	32	235	73	0	120
LED-C8	0	0	0	0	0	0	39	421

Table 3. Confusion matrix for CNN classifier for 8 instantiations of an LED light (Cree brand). The train and test samples were taken after 2 days of separation. Total accuracy of 76.11%.

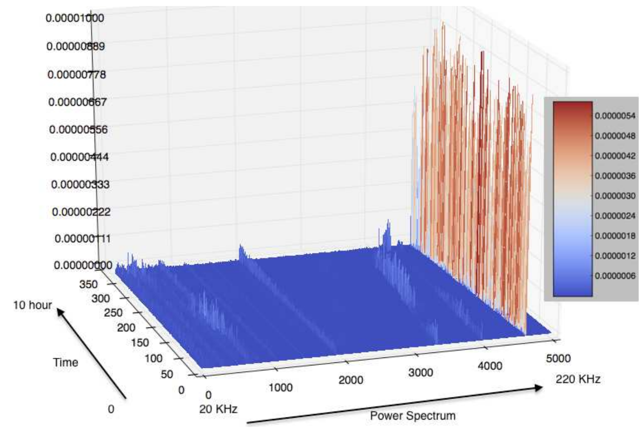


Fig. 4. A plot of the FFT features of an LED (Verbatim brand) over 10 hours. Within certain frequency bands, a spectral drift is observed.

4.3. Indoor Localization Using Light Fingerprints

Aside from the above controlled experiments with light bulbs, our light fingerprinting hardware of Figure 1 was tested in several indoor environments; e.g., hotel hallway, office building. The hardware system was able to distinguish multiple locations consistently based on CFL lights. Certain brands of LEDs (e.g., Verbatim) were challenging to distinguish consistently over days of time separation between train and test. One particular issue which causes errors in classification is plotted in Figure 4. The changes in the frequency magnitude values of one instantiation of an LED light bulb (Verbatim brand) are recorded over 10 hours. Within bands of frequency, spectral drift is observed over time. The drift regions may either be excluded as features, or compensated by giving more training data to the classifiers. A practical idea which improved classification accuracy for both CFLs and LEDs is the use of *multiple* feature vectors during the test phase of classifying a light bulb. In other words, during a *longer* test phase, a consensus may be reached over multiple feature vectors.

5. DISCUSSION

As shown by the classification results of experiments, it is possible to distinguish high-frequency fluctuations of different light bulbs, even of the same brand. Most CFL bulbs exhibit a greater variability in high-frequency switching patterns than LED bulbs. With more training data, the CNN classifier has the potential to perform better than the k-nearest neighbor classifier; however, classifiers with low-complexity for training and testing are good candidates for mobile devices. While a major property of our light fingerprinting system is its infrastructure-less ease of use, it is important to note that our system is still compatible with modulation systems which embed identifying signals in light switching patterns.

6. REFERENCES

[1] J. Vongkulbhisal, B. Chantaramolee, Y. Zhao, and W. S. Mohammed, "A fingerprinting-based indoor localization system using intensity modulation of light emitting diodes," *Microwave and Optical Technology Letters*, vol. 54, no. 5, pp. 1218–1227, 2012.

[2] G. Kail, P. Maechler, N. Preyss, and A. Burg, "Robust Asynchronous Indoor Localization Using LED Lighting," in *Proc. of IEEE Intern. Conf. On Acoustics, Speech And Signal Proc. (ICASSP)*, New York, 2014.

[3] K. Panta and J. Armstrong, "Indoor localisation using white LEDs," *Electronics Letters*, vol. 48, no. 4, pp. 228–230, Feb. 2012.

[4] M. Yasir, S. W. Ho, and B. N. Vellambi, "Indoor positioning system using visible light and accelerometer," *Journal of Lightwave Technology*, vol. 32, no. 19, pp. 3306–3316, Oct 2014.

[5] N. Rajagopal, P. Lazik, and A. Rowe, "Visual light landmarks for mobile devices," in *Proc. of the 13th Intern. Symp. on Inform. Proc. in Sensor Networks*, Piscataway, NJ, USA, 2014, IPSN '14, pp. 249–260, IEEE Press.

[6] Liqun Li, Pan Hu, Chunyi Peng, Guobin Shen, and Feng Zhao, "Epsilon: A visible light based positioning system," in *11th USENIX Symp. on Networked Systems Design and Implementation (NSDI)*, Seattle, WA, 2014, pp. 331–343, USENIX Association.

[7] Ye-Sheng Kuo, Pat Pannuto, Ko-Jen Hsiao, and Prabal Dutta, "Luxapose: Indoor positioning with mobile phones and visible light," in *Proc. of the 20th Annual Intern. Conf. on Mobile Computing and Networking*, New York, NY, USA, 2014, MobiCom '14, pp. 447–458, ACM.

[8] N. Ravi and L. Iftode, "FiatLux: Fingerprinting rooms using light intensity," in *Adj. Proc. of the Fifth Intern. Conf. on Pervasive Computing*, 2007.

[9] M. Azizyan, I. Constandache, and R. Roy Choudhury, "Surroundsense: Mobile phone localization via ambience fingerprinting," in *Proceedings of the 15th Annual International Conference on Mobile Computing and Networking*, New York, NY, USA, 2009, MobiCom '09, pp. 261–272, ACM.

[10] Q. Xu, R. Zheng, and S. Hranilovic, "Idyll: Indoor localization using inertial and light sensors on smartphones," in *Proc. of the 2015 ACM Intern. Joint Conf. on Pervasive and Ubiquitous Computing*, New York, NY, USA, 2015, UbiComp '15, pp. 307–318, ACM.

[11] Chi Zhang and Xinyu Zhang, "Litell: Robust indoor localization using unmodified light fixtures," in *Proceedings of the 22Nd Annual International Conference on Mobile Computing and Networking*, New York, NY, USA, 2016, MobiCom '16, pp. 230–242, ACM.

[12] Chi Zhang, Shipei Zhou, and Xinyu Zhang, "Visible light localization using incumbent light fixtures: Demo abstract," in *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM*. ACM, 2016, pp. 304–305.