Visible Light Networking and Sensing

Xia Zhou and Andrew T. Campbell
Department of Computer Science, Dartmouth College
{xia,campbell}@cs.dartmouth.edu

ABSTRACT
We propose for the first time an integrated Visible Light Communication (iVLC) system, which combines scalable VLC networking and accurate VLC sensing of mobile users. To meet this goal, we envision using modulated LED lights for communications between networked devices, while at the same time using the very same lights to accurately identify and track users, and importantly, sense and infer their gestures (e.g., pointing to an object in the room) as a means of collecting user analytics and enabling interactions with objects in smart spaces. Enabling the iVLC vision requires reliable VLC networking and robust VLC sensing. We discuss the key research components and open challenges in realizing this vision. By combining VLC networking and sensing, iVLC opens the way for a new class of context-aware applications and a new HCI paradigm not possible before.

Categories and Subject Descriptors
C.2.1 [Network Architecture and Design]: Wireless communication

Keywords
Visible light communication; Wireless networking; Sensing

1. INTRODUCTION
Mobile computing and communication renovation is accelerating. Within the next two years, the volume of data from mobile devices (e.g., smartphone) will surpass all data generated by wired networks [2]. The advent of new wearable devices (e.g., Google Glass, Samsung smart watch, Fitbit) and health sensors (e.g., biometric sensors printed onto the skin) is placing significant new demands on already limited radio spectrum. Today, people wear Google Glass, smart watches and fitness devices and carry smartphones. Ten years from now, people will likely wear or carry 10s of wireless devices and interact with the Internet and computing infrastructure (e.g., smart spaces) in markedly different ways. To stave off the spectrum crunch, there is a need for radically different thinking, otherwise the crunch will morph into a crisis.

To address these pressing problems, we envision a new integrated networking and sensing environment for smart spaces (e.g., a room, building, train) based on visible light communications (VLC) that will advance how we can communicate and interact (e.g., advancing HCI). Specifically, our vision is an integrated VLC networking and sensing for smart spaces (iVLC). We imagine that all future indoor spaces will take advantage of the ubiquity of lights as a medium that integrates communication and passive sensing. VLC embeds data into changes in light intensity with such a high frequency that human eyes cannot perceive light changes. Unlike traditional RF radio systems that require complex signal processing for modulating/demodulating signals, VLC uses low-cost, energy-efficient Light Emitting Diodes (LEDs) to transmit data: a VLC transmitter encodes data into light intensity changes, and the receiver recovers the data using a photodetector or light sensor to monitor light changes. VLC has a number of appealing properties. It is secure (i.e., does not penetrate walls, resisting eavesdropping), energy efficient, operates in 390THz – 800THz unregulated spectrum band, and introduces no radio interference to electromagnetic devices. Thus, it holds great potential to help solve the spectrum crunch problem.

To realize this vision, we propose to advance the state of the art in VLC in two new directions:

• VLC networking: We consider VLC networking rather than communications over a single VLC link – that is to say, we propose the coordination of large arrays of densely collocated VLC links. While others have considered VLC in outdoor vehicular networks [1, 20] and ad-hoc low-rate networks among LEDs [27], we are the first to propose VLC networking to provide seamless connectivity for mobile devices in indoor smart space environments.

• VLC sensing: We propose to use VLC not only for communications but also as a passive sensing medium. Imagine, as a person moves around in a smart space while communicating with their VLC-enabled devices (e.g., smart glasses, watches, phones, tablets), they can also continuously gesture and interact with the room and objects in the room (e.g., a wall mounted display, computers, doors, windows, coffee machine) driven by shadows on the floor created from multiple light sources hitting the person’s body as they gesture – this is akin to making shadow puppets on the floor.
communications using iVLC. Uplink communications could transmit data independently. We consider only downlink to its VLC modulation unit to modulate light signals and fetch data from the Internet. Each LED light is connected with Power Line Communication (PLC) technology \[15, 8\] and transmit the same data rate cell correctional light beam, and creates a coverage area called links in the room. Each LED light emits spotlight-like directional light sources. The key challenges we face are as follows.

1. **VLC Networking**

   The first key component of iVLC is VLC networking, which is to provide reliable, always-on VLC connectivity to mobile users across the complete coverage area (e.g., room, building, train). To achieve this goal, we need robust, reliable VLC networks. Existing work on VLC mostly focuses on improving performance of a single VLC link via advanced modulation schemes \[25, 9, 29, 30\]. Thus, we have to advance knowledge beyond a single VLC link to a network of VLC links. The key challenges we face are as follows.

   **Mobility Problem.** The atto-cell created by each LED light is very small in size (coverage area of 1-10m\(^2\)). Hence user mobility triggers frequent switches among atto-cells, resulting in connectivity losses and undesired latency. While

   \[1\] You might consider that this defeats the purpose of iVLC, but given the current energy constraint of wearables, supporting VLC full-duplex is very challenging. We consider VLC downlinks as the first step, and consider VLC uplinks as future work.
a considerable amount of work has looked at the handoff problem in cellular networks, we face new key challenges caused by atto-cell’s significantly smaller size and limited LED light sources.

To address the mobility problem, we propose to use predictive handoff schemes using real-time user tracking information (e.g. user location, moving direction, velocity) from VLC sensing. Based on user’s predicted movement, we involve the LED lights that the user is likely to encounter, minimizing the number of LED lights used for each user’s handoff while maintaining a smooth transition. Specifically, to track user mobility, we propose a shadow-based tracking method using the VLC sensing system. The key rationale is that the areas under a user’s feet are completely blocked by the user body and hence receive the lowest light illuminance intensity. Hence, by aggregating the illuminance data collected by light sensors on the floor, we can infer the user location by identifying the sensors that report the lowest illuminance data. Based on the estimated user locations over time, we can further infer user’s moving direction and velocity \( \vec{v} \). These user tracking information then serve as the input to the handoff scheme, allowing us to identify the potential cells that the user is likely to step into. These potential cells can begin transmitting the same data as the main cell, in anticipation of the user moving into the cell. We will examine schemes to determine the metric of picking handoff cells and the number of cells involved in handoff. We also plan to analyze metrics used for the mobile device to make the handoff decision.

**Interference Problem.** Visible light is diffusive in nature, and creates harmful interference to adjacent VLC receivers. Unlike typical wireless networks (e.g. Wi-Fi), an atto-cell network consists of links that cannot sense each other, because each transmitter is fixed on the ceiling and faces vertically downward to the floor. Hence the conventional carrier sensing scheme simply no longer applies. Furthermore, the receivers are wearable devices on each user’s body, and they can be worn at different positions (e.g. head, wrist, hand held). These contexts have different effects on the propagation of visible light, and hence complicate the interference condition.

To address the interference problem, we need to revisit interference modeling for atto-cell networks, and design mechanisms that separate interfering VLC transmissions in either the time or frequency domain. For interference modeling, we will examine both graph-based and SINR-based interference models. We need to take into account the 3D locations of wearable devices, and adapt existing interference models to the optical channel propagation, which is much more complicated than RF channels. Using the interference models, we will design scheduling algorithms that assign different time slots or frequency channels to interfering links. The scheduling algorithms are run in a centralized server connecting to all LED lights. We plan to explore TDMA scheduling algorithms for assigning time slots, and graph coloring algorithms for allocating frequency channels in our context.

We can also consider the dynamics of each atto-cell’s traffic demands, and design traffic-aware allocation algorithms to match the channel utilization to actual traffic loads.

**Network Deployment Problem.** Determining the optimal placement and density of LED lights in a smart space is non-trivial, because of the tradeoff between communicat-

4. **VLC SENSING**

The second key component of iVLC is robust VLC sensing, which uses VLC as a sensing medium where users can passively and unobtrusively interact with objects in smart spaces. Unlike prior work that uses cameras [22], sound waves [19], or Wi-Fi signals [3] to recognize gestures, the iVLC framework uses shadows to recognize gestures and track the mobility of users as they freely move around iVLC smart spaces. A key advantage of VLC sensing is that the intelligence needed to decode gestures is embedded in the communication medium (the light) and thus presents no overhead other than the need for passive light sensors deployed on the floor.

In brief, the VLC sensing system works as follows. Light sensors on the floor compute the illuminance of each light source using embedded beacons (e.g., \( l_1 \), \( l_2 \)). The composite shadow called the shadow map consists of one or more shadow components (Figure 3). The server connected to light sensors computes the shadow map and its shadow components. It then uses a gesture recognition classifier to infer user’s gesture (e.g., pointing in a direction). The system continuously tracks users and seeks gestures in real time. Because of the variability of how users gesture and the difference in room setups, we use an offline training to capture shadow maps for a set of distinctive gestures (e.g., point with hand, wave hand, push hand, pull hand, point with
both hands) for diverse users (e.g., body size, age, gender). Shadow components are used to build a mapping template for the gesture recognition system.

We next describe the key problems in VLC sensing and discuss possible solutions.

**Shadow Capture Problem.** Given multiple LED lights on the ceiling, the illuminance perceived by each light sensor represents a combination of lights from multiple LEDs from different locations. The light mixture diminishes the shadow caused by human body blockage, making shadow capture very challenging.

To address the shadow capture problem, we can leverage the fact that lights are actively communicating via VLC. We propose to embed a unique beacon into the light from each light source, creating a signature for light emitted from this light source. The beacons allow us to separate combined light signals and capture shadow components resulting from individual lights. These shadow components exhibit more distinguishable patterns and provide the basis for gesture recognition.

To demonstrate the feasibility of our idea, we have built a simulator that emulates the impact of human body blockage of light. The input of the simulator includes the room setting (e.g., room size, ceiling height), LED light and light sensor configurations (e.g., transmit power, Field-of-Vision), and the deployment (e.g., location, density) of LED lights and light sensors. The simulator emulates human body parts (e.g., arms, legs) using cylinders, follows the optical channel propagation model in [16], and computes the light illuminance perceived by light sensors on the floor in the presence of a user gesture. By default, we configure LEDs and light sensors as [16].

Using the simulator, Figure 4 compares the composite shadow map under two light sources $l_1$ and $l_2$, and the shadow components corresponding to individual light sources, when the user is making a pointing gesture. Clearly the shadow component of $l_1$ presents the most distinguishable pattern for recognizing this gesture. To realize the shadow separation, we need to study how to embed beacons into data transmissions with minimal overhead, how to leverage interference management scheme to avoid beacon collisions, and how to capture shadow when multiple users are present.

**Gesture Recognition Problem.** Different gestures can result in similar shadow maps even with a single light source, and the same gesture can result in slightly different shadow maps for different users. Therefore, it is complex to link an observed shadow map and its components to a gesture with confidence.

To distinguish gestures using shadow maps and shadow components, we can design machine-learning algorithms (e.g., k-NN) to identify the best reference light (or a set of reference lights) for each gesture so that the resulting shadow component(s) is the most distinguishable compared to their shadow components of other gestures. We will construct a dictionary of distinctive gestures (e.g., point with hand, wave hand, push hand), and use offline training to capture shadow components for a diverse group of users (e.g., gender, size) conducting these gestures. These shadow components are used to build a mapping template for the online gesture recognition system. We plan to evaluate image classification algorithms such as Random Forest for reliable, efficient gesture classification, and adaptive boost algorithm to improve the classification accuracy. Another interesting problem is to go beyond single-snapshot gestures, and to explore how to aggregate a sequence of gestures and identify user’s consecutive gestures (e.g., point and then wave the hand). This capability will open up new opportunities to transform the interactions between users and their environments.

**Sensing Data Sparsity Problem.** Limited by light sensor’s Field of Vision (FoV) and sensor density, light illumination data collected by light sensors is sparse in nature. This results in low-resolution shadow maps, making gesture recognition more difficult. Thus, we need to recover the full observation data from limited samples.

To address the sparsity problem of sensing data, we use the application of compressive sensing techniques to recover full observations from partial sensing data. We can adapt compressive sensing techniques to shadow capture and gesture recognition. We can also use computer vision algorithms to determine the optimal sensor locations and thus reduce the number of sensing data that need to be recovered.

## 5. RELATED WORK

**VLC Physical Layer.** The research group led by Nakagawa at Japan pioneered the early work on VLC [15, 19]. They envisioned the integration of VLC with the power-line communication (PLC) [15] to provide ubiquitous Internet access indoors, and first studied and simulated the propagation properties of VLC in indoor settings [16]. Since then, there have been tremendous efforts on boosting VLC data rate, by either filtering out the yellow light with a slow response time [17], or designing sophisticated modulation schemes (e.g., CSK [25], OFDM [8]), or using MIMO to enable parallel data streams [26]. Recent research breakthrough has achieved 10Gbps data rate using specialized LED lights consisting of arrays of micro-LEDs [11, 29]. To date, this is the fastest VLC system.

Advancements on VLC physical layer techniques are certainly encouraging. Our vision leverages and benefits from these physical layer techniques. While most existing work focuses on single-link performance in a controlled lab environment, our ideas go beyond a single VLC link, and study systems challenges in a VLC network.

**VLC Applications.** VLC application scenarios span from indoor communication [19], indoor localization [13, 15, 31], screen-camera communication [5, 10, 21, 24], vehicular networks [20, 24], underwater communication [20], to in-flight entertainment [21]. In particular, [20] is the first to build a VLC prototype for vehicular network and examine the networking properties in real outdoor systems. Our proposed research on VLC networking is inspired by this work, and focuses on understanding VLC networking performance in the indoor environment. In [5, 19], Ashok et al consider boosting the data rate of screen-camera links via the concept of visual MIMO. In [11], Hao et al design a new 2D barcode to address the screen size limitation on a smartphone, and the problem of image blur in mobile environment. In [13], Hu et al tackle the problem of imperfect frame synchronization. [27, 21] study LED-to-LED communication.

We are inspired by these VLC applications. We aim to broaden the use case of VLC, and explore the feasibility of using VLC to enable sensing applications and possibly a new paradigm of interactions between users and environments.
Sensing and Gesture Recognition. Active research examines user sensing and gesture recognition using cameras [23], sound waves [10], and wireless RF signals [23]. In [10], Gupta et al. leverage the Doppler effect of sound waves to identify user gesture. In [23], Pu et al. extend the idea to Wi-Fi signals. In [3], Adib and Katabi enable Wi-Fi based gesture recognition by leveraging MIMO techniques and treating a moving user as an antenna array. Together with [4], they are good examples of using wireless signals not only for connectivity, but also for sensing. Our vision is to extend the use of VLC to sensing.

6. CONCLUSION

We described our vision of iVLC, a framework integrating VLC networking and VLC sensing for smart spaces. iVLC reuses lights as both a communication and sensing medium, and thus enables users to communicate and interact with smart spaces in a radically different way. We discussed the key research challenges in realizing VLC networking and VLC sensing, and potential solutions. iVLC can have far-reaching impact on networking, sensing, computer vision, and HCI by examining their interplay. Our next step is to explore options for building an experimental platform for iVLC, which consists of LED transmitters used in [12] and light sensors embedded in smart fabric or existing smart devices (e.g., Android Nexus 5 phones, Pebble smart watches). We plan to implement and demonstrate the feasibility of our core ideas using the testbed.

While VLC is attracting significant interests in academia and industry, there is little known about the performance of VLC or how one would build or deploy VLC networks in the workspace or home. There is also little known about the types of applications VLC will enable. VLC has not been positioned as a mobile technology and its applicability and success in this domain is in question. By proposing iVLC, we aim to shine a light on these open issues. The proposed research will accelerate the deployment of VLC as the next-generation wireless communication to enable ubiquitous interactions between smart devices and smart spaces. It will advance our understanding of the necessary algorithmic advances needed to build VLC networks.

7. REFERENCES