Supplementary Material for:
Light Codes for Fast Two-Way Human-Centric Visual Communication

MOHIT GUPTA, University of Wisconsin–Madison, USA
JIAN WANG, KARL BAYER, and SHREE K. NAYAR, Snap Research, USA

In this supplementary report, we provide additional details supporting the content in our main article.

1 DETAILS OF MONTO CARLO SIMULATIONS
To validate the key theoretical result of the article (the optimal transmission strategy given in Equation (4)), we perform Monte Carlo simulations of the proposed stochastic coding scheme. We assigned random start times to transmission windows of both devices (to simulate lack of synchronization). For each device, the transmission window is divided into blocks; each device transmits during a block with probability \( p_t \). A device successfully transmits a code in a block if the other device is not transmitting during that time. Finally, a device decodes a received block successfully with a probability \( p_n \) (channel noise). For each choice of parameters \( (N, p_t, p_n) \), we compute the probability of successful transmission \( p_{\text{success}} \) by repeating the preceding process 100,000 times and counting the number of successes. The simulations were implemented in MATLAB. We will release the code upon the article’s acceptance.

2 STOCHASTIC TRANSMISSION PROTOCOLS IN COMMUNICATION LITERATURE
Stochasticity is utilized in several shared-medium access communication protocols. For example, ALOHA [Abramson 1985, 2009] (and its variants like slotted ALOHA [Martin 2005; Roberts 1973, 1975]) is classical and one of the most popular shared-medium communication protocols used in mobile wireless networks [Pahlavan and Levesque 1994; Stavenson 1984], wired (cable) networks such as Ethernet [Martin 2005; Metcalfe and Boggs 1976], and modern satellite networks [Abramson 1990].

Slotted ALOHA [Martin 2005; Roberts 1973, 1975] requires transmitters to transmit only in specified slots, and the slot boundaries are globally synchronized. This achieves higher efficiency and overall performance but requires global synchronization, which may not be possible in the uncontrolled user interface scenario considered here. Later, CSMA (carrier sensing medium access) [Bharghavan et al. 1994; Kleinrock and Tobagi 1975; Tanenbaum 2003] protocols were also introduced to improve the overall efficiency (by listening before sending to avoid packet collisions), at the cost of increased system complexity, and additional resources (power and time) for carrier sensing and coordination.

These classical approaches employ a “stochastic backoff” strategy, where a transmitter, in the case of a clash, waits for a random amount of time before re-transmitting. This approach requires carrier sensing (CSMA) [Kleinrock and Tobagi 1975; Metcalfe and Boggs 1976], and/or repeated acknowledgments from the receiver before making a decision to re-transmit, which could increase system complexity and latency, and thus is not ideal in a user interface setting.

3 DISCUSSION OF EXISTING HUMAN-CENTRIC COMMUNICATION TECHNIQUES
Wi-Fi. Radio wave based communication methods such as Wi-Fi can achieve a high data rate over long distances. However, Wi-Fi-based transmission and communication do not have directionality, resulting in a wide communication cone and low selectivity. Due to the lack of selectivity, Wi-Fi-based inter-device communication and pairing often requires manually selecting the device to communicate with, and perhaps even entering a password for authentication. A software protocol called Wi-Fi Direct (also known as peer-to-peer Wi-Fi) enables direct connection between two devices without the need of an access point. It inherits the high data rate and long-range benefits of Wi-Fi but also the limited selectivity, thus requiring manual selection and authentication. This reduces the overall fluidity of the user experience, which is critical in most human-centered consumer applications. Wi-Fi typically also has high power consumption (2–20 Watts), which is an important consideration in consumer mobile devices.

Bluetooth. Bluetooth is a wireless technology standard used for connecting two devices and exchanging data. In contrast to Wi-Fi, which builds asymmetrical client-server connections and where two devices need to connect to an access point, Bluetooth can directly build a symmetrical connection between two devices. It uses 2.4-GHz radio wave and has a range of about 10 m. Although Bluetooth has a lower data rate (≈ 1/10 of Wi-Fi) and smaller range, the power consumption is also considerably lower than that of Wi-Fi. Since Bluetooth is also based on radio waves, which do not have strong directionality, Bluetooth devices have a wide communication cone leading to limited selectivity and potential security risks such as cyber-flashing where a user can receive unwanted data.

Li-Fi. Li-Fi is a communication technology that uses laser beams or LED sources for transmitting information [Haugen et al. 1986; Tsonov et al. 2014], and high-speed photodiodes as receivers [Vucic et al. 2010]. These techniques, also referred to as visible light communication [Lee et al. 2007], although capable of achieving very high data rates in controlled settings, are not applicable in the human-to-human communication and consumer scenarios because they require near perfect alignment of the transmitter and the receiver. Instead of single photodiodes, visible light communication has also been explored by using high-speed CMOS cameras as receivers [Ashok et al. 2010; Matsushita et al. 2003;
Yuan et al. [2011], which do not easily lend themselves to most consumer devices due to high power and bandwidth requirements.

Receivers using hemispherical lenses [Barry and Kahn 1995] have been explored for non-directed Infrared (IR) communication systems [Otte et al. 2013] that do not require precise alignment between transmitter and receiver (e.g., television remote controls). These systems use IR protocols such as Sony SIRC [SONY 2022], the Phillips RC5 (Manchester encoding), and the NEC IR protocol (pulse distance encoding) [Digikey 2022]. These protocols and devices are typically meant for one-way communication; no handshake, authentication, or authorization takes place between the sender and the receiver. Various IR protocols, such as those based on pulse-position modulation (PPM) [Park and Barry 1995], have been proposed to improve the power efficiency and robustness to various non-idealities such as multi-path dispersion [Park and Barry 2004]. These and several other techniques in the IR communication literature [Ghassemlooy and Hayes 2000], although so far largely focusing on one-way communication, could be adapted to two-way interactions, thus further enhancing the performance of the proposed LICO system.

**Display-Camera Links.** A special case of Li-Fi is where the light source is a display. Data is transmitted typically by displaying a code (e.g., a QR code) on the display and capturing an image by a receiver camera. There have been several systems that use display-camera links for human-to-human communication [Hao et al. 2012; Perli et al. 2010; Tamang and Kim 2021; Wang et al. 2014]. Various codes have been considered, including frequency-domain spatial coding [Perli et al. 2010], color codes [Hao et al. 2012], spatial-temporal codes [Hu et al. 2014; Langlotz and Bimber 2007], and high-speed temporal codes with rolling shutter cameras [Jo et al. 2016; Woo et al. 2012]. The code could be displayed directly on the screen or encoded in temporal changes of the display’s alpha channel [Li et al. 2015]. An important consideration in uncontrolled human-to-human interaction is to handle the lack of synchronization between the display and camera devices [Hu et al. 2013; Langlotz and Bimber 2007]. Another kind of spatial code based modality was based on the bokeh effect, requiring cameras to capture a defocused image [Mohan et al. 2009]. Although display-camera links can provide interference-free directional communication in consumer devices, their range is limited due to limited camera resolution, and imaging degradations such as perspective distortion, motion blur, and ambient illumination [Perli et al. 2010]. Furthermore, display-camera-based interaction often requires users to clear their display (to pull up the code) and tilt their screen to the other user. These steps increase the overall interaction time. Although these limitations have been addressed to some extent by building robustness to common imaging degradations [Perli et al. 2010] and oblique displays [Jo et al. 2016; Wengrowski and Dana 2019], the tradeoff is relatively lower ergonomics and fluidity of user interactions.

An important class of display-camera communication methods are those based on steganography [Baluja 2017; Kessler and Hosmer 2011; Parikka 2017; Tancik et al. 2020; Zhang et al. 2020], where the goal is to make the displayed code imperceptible to humans [Jo et al. 2016; Nguyen et al. 2016; Tran et al. 2021; Wengrowski and Dana 2019; Yuan et al. 2012]. These methods allow displays to show image/video content while simultaneously transmitting imperceptible codes that can be received by cameras. In contrast, the goal of LICO is to develop a communication modality (and associated device) dedicated for intentional human-human interaction, which trades off the flexibility of steganography-based methods for fast and fluid human-centered communication.

**Near-Field Communication.** Near-Field Communication (NFC) is a set of protocols for very short distance (<2 cm) data transfer between two devices [Coskun et al. 2013; Paus 2007]. Similar to LICO, NFC methods are typically optimized for transferring a small amount of data, and are used for sharing small files or as the first step to build connection followed by a fast wireless methods for sharing large files. NFC works only for a very short range, thus requiring close contact and large interaction times for sharing information, which may not be desirable in human-centric application scenarios. In contrast, the proposed system based on Light Codes has a narrow field of view (whereas NFC has a narrow distance range), and enables fast communication and fluid user interactions across a range of distances, while maintaining low power requirements and minimal security risks.

**Contact-Sharing Devices.** There are dedicated hardware devices called digital business cards that users can buy off-the-shelf for contact sharing. A user can configure one with their contact information and social media accounts, and another user can use a cellphone to scan it to get such information. These devices are made in the form of a coin-sized sticker tag [Dot 2022], a key fob [Popl 2020], a wristband [Popl 2022], a watch band [Linq 2022], a hub [Linq 2022c], a badge [Linq 2022a] or a business card sized plastic [Social Master 2022] or metal card [V1CE 2022], and use technology of NFC or QR code. Some of these are also available in the cellphone case form factor, similar to the proposed LICO devices.

**REFERENCES**


