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_{suc} 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; Stavenow 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; Tsonev 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;

0730-0301/2023/09-ART1 $15.00
https://doi.org/10.1145/3617682

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 de-focused 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 ergonomy 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 2022a], a wristband [Popl 2022b], a watch band [Linq 2022b], 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

Norman Abramson. 1985. Development of the ALOHANET. *IEEE Transactions on Information Theory* 31, 2 (1985), 119–123.

Norman Abramson. 1990. VSAT data networks. *Proceedings of the IEEE* 78, 7 (1990), 1267–1274.

Norman Abramson. 2009. The ALOHAnet—Surfing for wireless data. *IEEE Communications Magazine* 47, 12 (2009), 21–25.

Ashwin Ashok, Marco Gruteser, Narayan Mandayam, Jayant Silva, Michael Varga, and Kristin Dana. 2010. Challenge: Mobile optical networks through visual MIMO. In *Proceedings of the 16th Annual International Conference on Mobile Computing and Networking (MobiCom'10).* 105–112.

Shumeet Baluja. 2017. Hiding images in plain sight: Deep steganography. In *Advances in Neural Information Processing Systems*, I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Eds.), Vol. 30. Curran Associates, 1–11. https://proceedings.neurips.cc/paper/2017/file/838e8afb1ca34354ac209f53d90c3a43-Paper.pdf

John R. Barry and Joseph M. Kahn. 1995. Link design for nondirected wireless infrared communications. *Applied Optics* 34, 19 (July 1995), 3764–3776. https://doi.org/10.1364/AO.34.003764

Vaduvur Bharghavan, Alan John Demers, Scott J. Shenker, and Lixia Zhang. 1994. MACAW: A media access protocol for wireless LANs. *ACM SIGCOMM Computer Communication Review* 24, 4 (1994), 212–225.

Vedat Coskun, Busra Ozdenizci, and Kerem Ok. 2013. A survey on near field communication (NFC) technology. *Wireless Personal Communications* 71, 3 (2013), 2259–2294.

DigiKey. 2022. Understanding the Basics of Infrared Communications. Retrieved November 1, 2022 from https://www.digikey.com/en/maker/blogs/2021/understanding-the-basics-of-infrared-communications

Dot. 2022. Shop Basics. Retrieved August 1, 2023 from https://dotcards.net/collections/all-products

Z. Ghassemlooy and A. R. Hayes. 2000. Digital pulse interval modulation for IR communication systemsâĂŤA review. *International Journal of Communication Systems* 13, 7-8 (2000), 519–536.

Tian Hao, Ruogu Zhou, and Guoliang Xing. 2012. COBRA: Color barcode streaming for smartphone systems. In *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services (MobiSys'12)*. 85–98.

P. R. Haugen, S. Rychnovsky, A. Husain, and L. Hutcheson. 1986. Optical interconnects for high speed computing. *Optical Engineering* 25, 10 (1986), 251076.

Wenjun Hu, Hao Gu, and Qifan Pu. 2013. LightSync: Unsynchronized visual communication over screen-camera links. In *Proceedings of the 19th Annual International Conference on Mobile Computing and Networking (MobiCom'13)*. 15–26.

Wenjun Hu, Jingshu Mao, Zihui Huang, Yiqing Xue, Junfeng She, Kaigui Bian, and Guobin Shen. 2014. Strata: Layered coding for scalable visual communication. In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking (MobiCom'14)*. ACM, New York, NY, 79–90.

Kensei Jo, Mohit Gupta, and Shree K. Nayar. 2016. DisCo: Display-camera communication using rolling shutter sensors. *ACM Transactions on Graphics* 35, 5 (2016), Article 150, 13 pages.

Gary C. Kessler and Chet Hosmer. 2011. An overview of steganography. In *Advances in Computers*, Vol. 83. Elsevier, 51–107.

Len Kleinrock and Fouad A. Tobagi. 1975. Packet switching in radio channels: Part I—Carrier sense multiple-access modes and their throughput-delay characteristics. *IEEE Transactions on Communications* 23, 2 (1975), 1400–1416.

Tobias Langlotz and Oliver Bimber. 2007. Unsynchronized 4D barcodes: Coding and decoding time-multiplexed 2D colorcodes. In *Proceedings of the International Conference on Advances in Visual Computing*.

C. G. Lee, C. S. Park, J. H. Kim, and D. H. Kim. 2007. Experimental verification of optical wireless communication link using high-brightness illumination light-emitting diodes. *Optical Engineering* 46, 12 (2007), 125005.

Tianxing Li, Chuankai An, Xinran Xiao, Andrew T. Campbell, and Xia Zhou. 2015. Real-time screen-camera communication behind any scene. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys'15)*. 197–211.

Linq. 2022a. Linq Badge. Retrieved January 1, 2022 from https://buy.linqapp.com/collections/all-1/products/linq-badge

Linq. 2022b. Linq Band for Apple Watch. Retrieved January 1, 2022 from https://buy.linqapp.com/products/linq-band-for-apple-watch

Linq. 2022c. Linq Hub. Retrieved January 1, 2022 from https://buy.linqapp.com/products/linq-hub

J. Martin. 2005. The impact of the DOCSIS 1.1/2.0 MAC protocol on TCP. In *Proceedings of the 2nd IEEE Conference on Consumer Communications and Networking*.

N. Matsushita, D. Hihara, T. Ushiro, S. Yoshimura, J. Rekimoto, and Y. Yamamoto. 2003. ID CAM: A smart camera for scene capturing and ID recognition. In *Proceedings of the 2003 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'03)*.

Robert M. Metcalfe and David R. Boggs. 1976. Ethernet: Distributed packet switching for local computer networks. *Communications of the ACM* 19, 7 (July 1976), 395âĂŞ404.

Ankit Mohan, Grace Woo, Shinsaku Hiura, Quinn Smithwick, and Ramesh Raskar. 2009. Bokode: Imperceptible visual tags for camera based interaction from a distance. *ACM Transactions on Graphics* 28, 3 (2009), Article 98, 8 pages.

Viet Nguyen, Yaqin Tang, Ashwin Ashok, Marco Gruteser, Kristin Dana, Wenjun Hu, Eric Wengrowski, and Narayan Mandayam. 2016. High-rate flicker-free screen-camera communication with spatially adaptive embedding. In *Proceedings of the 35th Annual International Conference on Computer Communictions (IEEE INFOCOM'16)*. 1–9. https://doi.org/10.1109/INFOCOM.2016.7524512

Rob Otte, Leo P. De Jong, and Arthur H. M. Van Roermund. 2013. *Low-Power Wireless Infrared Communications*. Springer.

K. Pahlavan and A. Levesque. 1994. Wireless data communications. *Proceedings of the IEEE* 82, 9 (1994), 1398–1430.

Jussi Parikka. 2017. Hidden in Plain Sight: The Stagnographic Image. Retrieved January 1, 2022 from https://unthinking.photography/themes/fauxtography/hidden-in-plain-sight-the-steganographic-image

Hyuncheol Park and J. R. Barry. 1995. Modulation analysis for wireless infrared communications. In *Proceedings of the IEEE International Conference on Communications (ICC'95)*, Vol. 2. 1182–1186. https://doi.org/10.1109/ICC.1995.524287

Hyuncheol Park and J. R. Barry. 2004. Trellis-coded multiple-pulse-position modulation for wireless infrared communications. *IEEE Transactions on Communications* 52, 4 (2004), 643–651. https://doi.org/10.1109/TCOMM.2004.826382

Annika Paus. 2007. *Near Field Communication in Cell Phones*. Seminararbeit Ruhr-Universitat Bochum.

Samuel David Perli, Nabeel Ahmed, and Dina Katabi. 2010. PixNet: Interference-free wireless links using LCD-camera pairs. In *Proceedings of the 16th Annual International Conference on Mobile Computing and Networking (MobiCom'10)*. 137–148.

Popl. 2022a. Digital Business Card–Keychain. Retrieved January 1, 2022 from https://popl.co/collections/all-products/products/popl-starter-pack

Popl. 2022b. Digital Business Card–NFC Bracelet. Retrieved January 1, 2022 from https://popl.co/collections/best-sellers/products/popl-band-bogo

L. G. Roberts. 1973. Capture effects on ALOHA channels. In *Proceedings of the 6th Hawaii Conference on System Sciences*.

L. G. Roberts. 1975. ALOHA packet system with and without slots and capture. *ACM SIGCOMM Computer Communication Review* 5, 2 (1975), 28–42.

Social Master. 2022. Plastic NFC Digital Business Card. Retrieved January 1, 2022 from https://socialmaster.com/products/white-plastic-bc

Sony. 2022. Sony SIRC Protocol. Retrieved November 1, 2022 from https://www.sbprojects.net/knowledge/ir/sirc.php

B. Stavenow. 1984. Throughput-delay characteristics and stability considerations of the access channel in a mobile telephone system. *ACM SIGMETRICS Performance Evaluation Review* 12, 3 (1984), 105–112.

Lakpa Dorje Tamang and Byung Wook Kim. 2021. Deep D2C-Net: Deep learning-based display-to-camera communications. *Optics Express* 29, 8 (2021), 11494–11511.

Matthew Tancik, Ben Mildenhall, and Ren Ng. 2020. StegaStamp: Invisible hyperlinks in physical photographs. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR'20)*. 1515–1524. https://doi.org/10.1109/CVPR.2019.00161

A. Tanenbaum. 2003. *Computer Networks*. Prentice Hall.

Vu Tran, Gihan Jayatilaka, Ashwin Ashok, and Archan Misra. 2021. DeepLight: Robust & unobtrusive real-time screen-camera communication for real-world displays. In *Proceedings of the International Conference on Information Processing in Sensor Networks*.

Dobroslav Tsonev, Stefan Videv, and Harald Haas. 2014. Light fidelity (Li-Fi): Towards all-optical networking. In *Broadband Access Communication Technologies VIII*, Vol. 9007. International Society for Optics and Photonics, 900702.

V1CE. 2022. Metal Hybrid One Card. Retrieved January 1, 2022 from https://v1ce.co/products/metal-hybrid-one-card-designer

J. Vucic, C. Kottke, S. Nerreter, K.-D. Langer, and J. W. Walewski. 2010. 513 Mbit/s visible light communications link based on DMT-modulation of a white LED. *Journal of Lightwave Technology* 28, 24 (2010), 3512–3518.

Anran Wang, Shuai Ma, Chunming Hu, Jinpeng Huai, Chunyi Peng, and Guobin Shen. 2014. Enhancing reliability to boost the throughput over screen-camera links. In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking (MobiCom'14)*. 41–52.

Eric Wengrowski and Kristin Dana. 2019. Light field messaging with deep photographic steganography. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR'19)*. 1515–1524. https://doi.org/10.1109/CVPR.2019.00161

Grace Woo, Andy Lippman, and Ramesh Raskar. 2012. VRCodes: Unobtrusive and active visual codes for interaction by exploiting rolling shutter. In *Proceedings of the 2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR'12)*.

Wenjia Yuan, Kristin Dana, Ashwin Ashok, Marco Gruteser, and Narayan Mandayam. 2012. Dynamic and invisible messaging for visual MIMO. In *Proceedings of the IEEE Workshop on the Applications of Computer Vision (WACV'12)*. 345–352. https://doi.org/10.1109/WACV.2012.6162992

W. Yuan, K. Dana, M. Varga, A. Ashok, M. Gruteser, and N. Mandayam. 2011. Computer vision methods for visual MIMO optical systems. In *Proceedings of the IEEE International Workshop on Projector-Camera Systems (Held with CVPR)*.

Chaoning Zhang, Philipp Benz, Adil Karjauv, Geng Sun, and In So Kweon. 2020. UDH: Universal deep hiding for steganography, watermarking, and light field messaging. In *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin (Eds.), Vol. 33. Curran Associates, 10223–10234. https://proceedings.neurips.cc/paper/2020/file/73d02e4344f71a0b0d51a925246990e7-Paper.pdf