When Visible Light (Backscatter) Communication Meets Neuromorphic Cameras in V2X

Kenuo Xu[P], Kexing Zhou[P], Chengxuan Zhu[P], Shanghang Zhang[P], Boxin Shi[P]
Xiaoqiang Li[PS], Tiejun Huang[P], Chenren Xu[PZK✉*]

[P]Peking University [S]SpikeVision (Beijing) Technology Co., Ltd. [Z]Zhongguancun Laboratory
[K]Key Laboratory of High Confidence Software Technologies, Ministry of Education (PKU)

ABSTRACT

Intelligent transportation systems are predicted to change the way people live in the foreseeable future. Vehicular networks are one of the key enablers for such systems, yet no status-quo solutions of vehicular networks make practical deployments possible. This paper proposes NeuromorphicVLC, a visible light communication system equipped with neuromorphic cameras as optical receivers to improve its performance. Compared with conventional photodiodes or cameras, the new type of bio-inspired CMOS vision sensors highlight high temporal resolution, large dynamic range, and adequate spatial resolution to filter out ambient noise. We develop a complete signal processing pipeline to detect the VLC transmitters and demodulate the messages. Preliminary experimental results demonstrate NeuromorphicVLC achieves a 4.8 Kbps bit rate and ensures reliability in various range and mobile scenarios.

CCS CONCEPTS

• **Networks** → **Mobile networks**; • **Hardware** → *Signal processing systems*; • **Computing methodologies** → *Computer vision.*

KEYWORDS

Visible Light Backscatter Communication; V2X; Neuromorphic Camera; Spike Camera

ACM Reference Format:
Kenuo Xu, Kexing Zhou, Chengxuan Zhu, Shanghang Zhang, Boxin Shi, Xiaoqiang Li, Tiejun Huang, Chenren Xu. 2023. When Visible Light (Backscatter) Communication Meets Neuromorphic Cameras in V2X . In *The 24th International Workshop on Mobile Computing Systems and Applications (HotMobile '23), February 22–23, 2023, Newport Beach, CA, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3572864.3580333

1 INTRODUCTION

Transportation systems are becoming ever closer to a major technological transformation. Vehicles, which used to be bare metals-on-wheels, have been equipped with hundreds of sensors and embedded computers to assist human drivers with (semi-)autonomous driving capabilities. Along with this trend, the road infrastructure is also evolving such as adaptive traffic lights and networked pay tolls. An essential and anticipated component for the forthcoming intelligent transportation system is Vehicle-to-Everything (V2X) communication that enables effective information exchange between the vehicles and the infrastructure.

The majority of V2X systems proposed until now are based on Radio Frequency (RF) communication, such as the IEEE 802.11p protocol [1] based on WLAN and C-V2X solutions [2] upon cellular networks. Although such technologies have been well investigated, a few fatal factors hinder the RF-based V2X solutions from scalable applications, such as the scarce radio spectrum and severe inter-tag interference in scenarios with high node density [3]. Such fundamental restrictions of RF prompt a complementary, if not alternative, V2X communication technology based on a different physical medium. Visible light, when utilized as a wireless communication medium, can potentially tackle the imperfections of RF. In recent years, Visible Light Communication (VLC) has been regarded as a promising technology for the ubiquitous 6G wireless networks [4, 5]. Benefiting from the massive bandwidth in the electromagnetic spectrum and line-of-sight propagation nature of visible light, VLC is envisioned to complement radio frequency that has several physical drawbacks such as the spectrum crunch problem [3]. In addition to the comprehensive academic research progress on indoor (short-range) VLC systems [6], VLC has been deployed in several practical real-world deployments such as intra-cabin communication for business jets [7], and has been standardized by IEEE [8] and ITU [9]. Moreover, the application scenarios of VLC have also been expanded to outdoor vehicular scenarios [10].

Despite the nascent research progress of vehicular VLC, innovative solutions are needed to support complex on-road scenarios. In order to receive the optical links in the presence of high mobility and adverse lighting conditions, a vehicular VLC receiver should meet three practical requirements: *high sampling rate* (to enable a data rate higher than Kbps level), *high dynamic range* (to handle sudden changes of lighting conditions), and *adequate resolution* (to exclude ambient optical noise from the area-of-interest). Status-quo VLC systems typically use *photodiodes* (with very few pixels for reception) as the optical receiver. However, the optical noise in the outdoor mobile scenarios is considerably more severe and leads to significant throughput/reliability reductions due to the insufficiency of resolution [11]. To handle the interference from ambient noise, researchers exploit *standard cameras* (with rolling-shutter effect) with a high number (*e.g.,* 1920 × 1080) of photodetectors to omit the undesirable optical signals within its field-of-view [12].

*K.Xu, K.Zhou and C.Zhu are the co-primary student authors.
✉: chenren@pku.edu.cn

HotMobile '23, February 22–23, 2023, Newport Beach, CA, USA

ACM ISBN 979-8-4007-0017-0/23/02…$15.00
https://doi.org/10.1145/3572864.3580333

Receiver	Sampling rate	Dynamic range	Resolution
PD	• 1 MHz	• 100 dB	1×1 for each
SC	60 Hz	60 dB	• 1920×1080
NC	• 40 kHz	• 120 dB	• 400×250

Table 1: Comparison of neuromorphic camera and related receivers. • means the property is adequate for vehicular VLC. PD: photodiode (array); SC: standard camera; NC: neuromorphic camera.

The main limitation originates from the low sampling rate of cameras – a typical camera captures only 30-60 frames per second, which provides an insufficient communication bandwidth (tens of bps). The rolling-shutter effect may enhance the data rate by tens or even hundreds of times [13]. However, the transmitter should be quasi-static to take full advantage of the rolling shutters. The deficiencies of conventional receivers stimulate research on specialized devices. High-speed cameras [14], digital micro-mirror devices [15], and dedicated CMOS sensors [16] have been explored as potential solutions, yet they are in the proof-of-concept stage with almost no real-world deployment. To sum up, no optical receivers meet the practical requirements of vehicular VLC so far, and innovative receivers would be necessary if vehicular VLC is to one day support practical V2X applications.

In this paper, we introduce *neuromorphic cameras* as the optical receivers in pursuit of a practical vehicular VLC system. Neuromorphic cameras, which mimic the neurobiological structures and functionalities of the biological retinas, provide a refreshing and promising perspective compared with conventional cameras [17]. They have a couple of advantages including high temporal resolution (in the order of microseconds) and high dynamic range (120 dB versus 60 dB of standard cameras) [18]. Such properties make them an appealing solution that allows high mobility and complex ambient light conditions for the receiver of vehicular VLC. The academia has noticed the emerging devices and started to apply them for VLC recently [19, 20], but still in their infant stages.

We propose NeuromorphicVLC, an hardware and software prototype of a VLC system equipped with neuromorphic cameras that aims at vehicular communication and networking. We select the bio-inspired spike camera [21] as the neuromorphic camera receiver, considering its advantage of high dynamic range (> 100 dB) and high sampling rate (up to 40000 Hz). Compared with other typical neuromorphic cameras such as event cameras [18], the spike camera can capture the visual images (from the neuromorphic output of the camera, *i.e.,* the spike streams) as a conventional camera without motion mismatch [22]. Tab. 1 shows the advantage of our receiver selection compared with other mature VLC receiver solutions. The transmitter is built upon the Visible Light Backscatter Communication (VLBC) technology [23] that features low power consumption (sub-mW, 100x less than typical VLC transmitters) and simplicity of network stack to implement the proposed approach. The physical and link-layer of VLBC have verified their feasibility in sparse vehicular communication [24].

We verify the design of NeuromorphicVLC primarily with in-lab experiments. Evaluation results show that NeuromorphicVLC achieves a near-zero bit error rate when the tags transmit at 4.8 Kbps. The links are robust under various distances and mobile scenarios. Our preliminary results demonstrate the feasibility of enhancing vehicular VLC with neuromorphic cameras. It is possible and promising to leverage the newly-invented camera technology as a desirable optical receiver for future V2X communication and networking systems. These results have important implications that may pose exciting new opportunities for the renovation of intelligent transportation systems.

2 BACKGROUND

2.1 Vehicular Visible Light Communication

Visible Light Communication (VLC) is regarded as a promising solution for vehicular networks because of its low cost and no electromagnetic interference. Modern illumination and display technologies, such as light-emitting diode (LED) and liquid-crystal display (LCD), are becoming more accessible and have enabled new perspectives for communication – fast and reliable networks can be accessed via the ubiquitous illumination infrastructure. Current VLC systems typically employ LEDs as the transmitter to modulate message bits (*e.g.,* with on-off keying) on the light signals. The optical signals are further received by certain receivers such as photodiodes [25, 26], (high-speed) cameras [13, 14], and dedicated sensors [15, 16] as the receiver. However, no status-quo VLC systems satisfy the practical requirements for vehicular networking. Photodiode-based solutions suffer from ambient interference (sunlight and other luminaries) greatly, or get restricted by a narrow field-of-view (FOV) [27]. Meanwhile, camera-based systems offer a wider FOV but are limited in the low throughput due to low temporal resolution (typically 30-60 frames per second). Novel devices thus become a necessity for practical VLC systems.

2.2 Spike Camera

The bio-inspired spike camera [21] is a representative of neuromorphic cameras that we mainly focus on in this paper. The spike camera, aiming to capture high-speed movements, is inspired by the sampling mechanism of primate fovea consisting of three main modules: a photoreceptor, an integrator, and a threshold comparator. The integrator keeps integrating the incident light intensity I. Once the voltage reaches a predefined threshold Φ, *i.e.,* $\int I dt \geq \Phi$, the comparator outputs a one-bit spike immediately and resets the integrator at the same time. In this way, the output of a spike camera is a 0-1 array (where 1 represents a spike and 0 is recorded for timing) as the sequence of *spike stream* for each independent pixel. The spike streams are bio-inspired representations of visual images. Specialized algorithms are required to perform vision tasks and extract the visual information from the spike streams, such as for visual reconstruction [22], optical flow estimation [28], depth estimation [29], and super-resolution [30]. The spike camera has several merits compared with conventional cameras, such as high temporal resolution (up to 40000 Hz), and high dynamic range (> 100 dB). These merits allow practical high-speed vision for autonomous driving, robotics, and unmanned aerial vehicles. In addition to the performance merits, the image sensor chip of the spike camera uses the common CMOS technology (*e.g.,* 110 nm) that has been widely used in conventional image sensors for smartphones. Therefore, the cost of a spike camera could be comparable to existing vision sensors once in massive manufacturing.

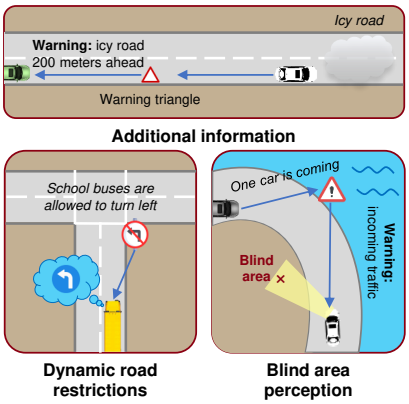


Figure 1: Application scenarios.

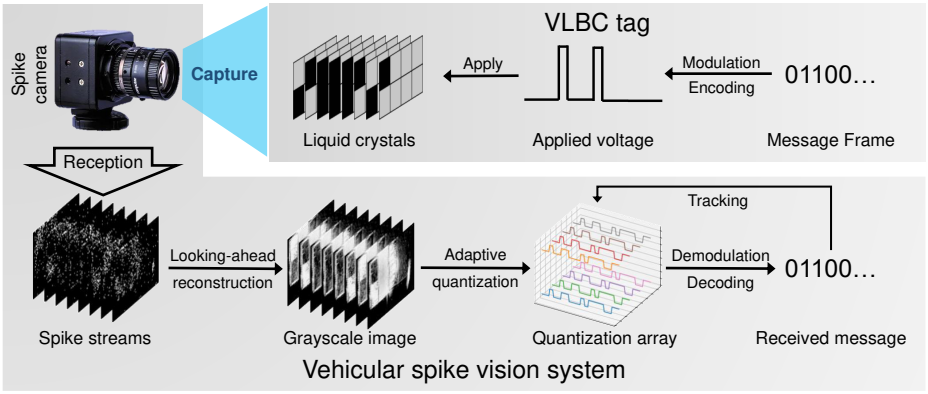


Figure 2: System architecture of NeuromorphicVLC.

2.3 Visible Light Backscatter Communication

Visible Light Backscatter Communication (VLBC) is an emerging technology that utilizes special low-power tags as the transmitters of VLC to reduce power consumption. The basic design of a VLBC tag is to reflect the incoming light using a retroreflective fabric and modulate it with a liquid crystal shutter. The retroreflective fabric, which has been widely used in on-road objects including road signs and safety clothing, reflects light to its source with minimum scattering. A liquid crystal shutter leverages the physical properties of liquid crystals to electronically control the illumination level of light passing through it. At the upper level, the communication logic flow is controlled by a low-power micro-control unit inside the tag. Status-quo VLBC tag transmits at as fast as 1 Kbps with less than 500 μW power consumption. Therefore, it can be powered by a small solar panel, or by a button cell with years of battery life [23]. In addition, the multi-pixel design, which allows the tag to control different parts of the liquid crystal shutters independently, can further improve the data rate to up to 32 Kbps [31]. At the receiver side, both (multi-)photodiodes [32] and smartphone cameras [33] have been explored to receive the modulated bits sent by tags.

3 NEUROMORPHICVLC DESIGN

3.1 System Overview

NeuromorphicVLC is designed to benefit V2X applications scenarios including providing additional information to the out-of-sight vehicles, broadcasting dynamic road restrictions (*e.g.,* bus-only lanes), and assisting perception in blind areas due to blockage (*e.g.,* in mountainous areas), as shown in Fig. 1. The VLBC tags are deployed near the road and can be integrated into the status-quo infrastructure such as road signs and warning triangles. They are equipped with liquid crystal shutters that modulate messages with optical signals according to the control of a micro-control unit. They can either be set up to broadcast (pre-loaded) static messages repeatedly or dynamic environmental sensory data. At the other end, the vehicles are equipped with spike cameras (*e.g.,* integrated with Vehicle Traveling Data Recorder), which may also be leveraged for visual perception tasks, to receive the messages sent by the tags. The spike streams captured by the spike camera are further processed with on-vehicle computers of the spike vision system.

We show the system architecture of NeuromorphicVLC in Fig. 2. The spike vision systems mount on the vehicles process the spike streams (the output of the spike camera) and demodulate the information sent by the tags with a series of algorithms. The algorithm framework is divided into four steps. First, we design an image reconstruction algorithm for spike cameras to reconstruct grayscale images from spike streams (§3.2). The algorithm limits the look-ahead depth of the spike streams, focusing on the bright pixels that are more likely to be the target. An adaptive quantization algorithm is then applied to the reconstructed grayscale images that can cope with different luminance for contrast enhancement (§3.3). We further design a demodulation algorithm, which leverages the physical property of the liquid crystal shutters, to extract the bit streams from the quantized images (§3.4). During the demodulation process, we also estimate the signal quality (§3.5) for each pixel and then choose the pixel with the best signal quality as the key point. Since the pixels on a particular control unit of the tag convey the same information, focusing on the key point allows low-error-rate demodulation. The key point is updated dynamically according to signal quality to track the motion of tags (§3.6).

3.2 Looking-ahead Image Reconstruction

An image reconstruction algorithm converts the spike streams into intensity-based images that could be analogized to sampling sequences in each pixel. We design an image reconstruction algorithm based on the texture from inter-spike interval [34]. We refine the algorithm with the following observation: spike cameras cannot achieve high temporal resolution because of the low spiking rate in dark areas. Therefore, we focus on the bright pixels for further processing, whereas the dark areas can be discarded for efficiency. We only deal with the pixels where the spike interval δ stays below a given threshold K, *i.e.,* our look-ahead depth is restricted to no more than K in the spike streams. Grayscale transformation is further used to increase the contrast of the reconstructed image, which can be expressed as:

$$lightness = \begin{cases} f(\delta/K), & \delta < K \\ f(1), & otherwise \end{cases}$$

f is the kernel of the grayscale transformation. We use three transformation kernels in our design, which are: $f_1(x) = \log\left(\frac{1}{K} + x\right)$, $f_2(x) = \sqrt{x}$, and $f_3(x) = x$. $1/K$ is added to $f_1(x)$ to avoid $\log(0)$.

3.3 Adaptive Quantization

The quantization algorithm detects the bright pixels and dark pixels in the transmitter's zone and quantizes them to +1 and −1 respectively, corresponding to the non-return-to-zero (NRZ) coding we use. Other pixels (noise) in the environment are set to 0. The idea is to calculate a shot-time upper bound *upper* and lower bound *lower* of luminance for each pixel. We use a derived threshold $(upper + lower)/2$ to quantize the signal to +1 or −1. If *upper* and *lower* are too close to each other (determined by another threshold *thres*), the pixel is recognized as noise and set to 0. In order to adapt to changes in brightness, *upper* and *lower* are adjusted over time to the same brightness as the signal. Each pixel is quantized to match the NRZ coding independently. Our design adapts to different pixels when there are multiple tags (each with different brightness) in the scene, or different pixels within a particular tag where the illumination is not exactly the same.

3.4 Demodulation

It is challenging to demodulate the signals sent from the VLBC tags because of the physical properties of the liquid crystal shutters: the gradual falling edges last more than 4 ms but sharp rising edges are less than 1 ms [31]. Our demodulation algorithm maintains *next* as the next sampling point and does clock recovery when the level changes. When the quantized signal changes, *next* is updated to $i + \frac{1}{3} \cdot period$ in the case of rising edge, and $i + \frac{2}{3} \cdot period$ for falling edge to treat the asymmetric edges differently. The algorithm outputs two arrays: *demoded* and *key_point*. *demoded* stores the most recent data captured on the pixel. Meanwhile, *key_point*, as well as *demoded*, are further to be leveraged to calculate the signal quality for tracking the key point.

3.5 Signal Quality Estimation

Although one tag in the field of view consists of multiple pixels, a lot of pixels are likely to be interrupted by object mobility and/or environmental noise. In addition, the grayscale images (output from §3.2) might also be noisy due to the random spikes from circuit noise. Therefore, it is crucial to find a good sampling position (*i.e.,* *key point*) with high signal quality to ensure the robustness of the demodulation process. We define signal quality as the sum of two parts, namely signal similarity and signal luminance. Thus, the signal quality of the pixel at (x, y) is:

$$Quality[x, y] = \alpha \cdot Similarity[x, y] + \beta \cdot Luminance[x, y]$$

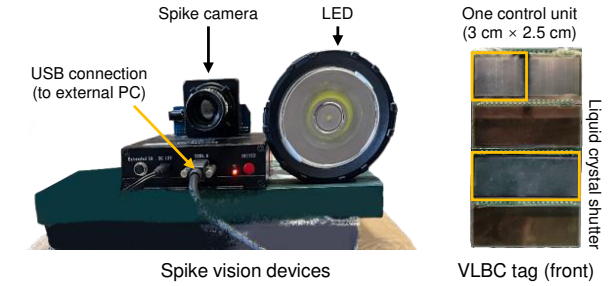
We measure the similarity by finding the number of identical signals demodulated in the neighborhood, and the luminance by *upper* obtained during adaptive quantization.

$$Luminance[x, y] = \frac{1}{max\_luminance} upper[x, y]$$

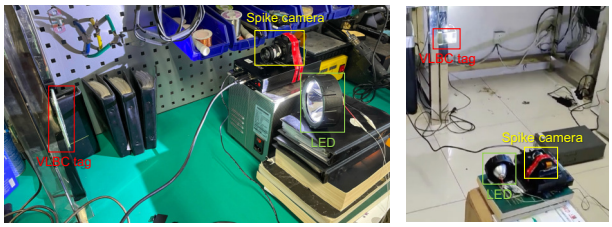
$$Similarity[x, y] = \frac{1}{(2N+1)^2} \sum_{i=x-N}^{x+N} \sum_{j=x-N}^{x+N} Sim(i, j, x, y)$$

$$Sim(i, j, x, y) = \left(demoded[i, j], demoded[x, y]\right)_M$$

The value of $(\cdot, \cdot)_M$ is 1 when the last M bits of the two operands are the same, and 0 otherwise; *max_luminance* is the maximum possible luminance.



(a) Devices.



(b) Static scenarios. **(c) Mobile scenarios.**

Figure 3: Experimental setup.

3.6 Key Point Tracking

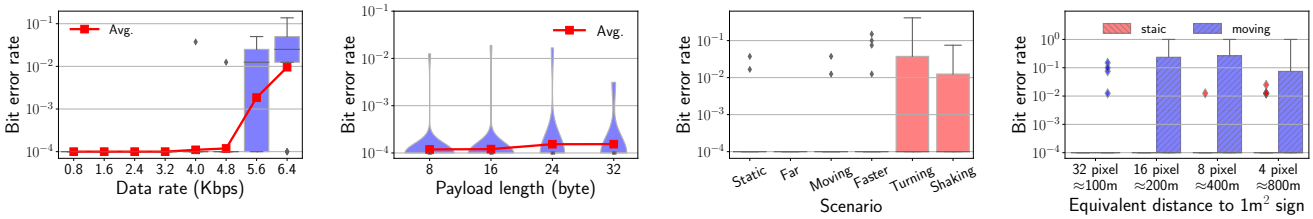
The output array *key_point* of §3.4 stands for the key points, *i.e.,* the pixels with the best signal quality among other pixels with the same device ID. The key point is generally at the center of the shutter, where the signal is the most stable and strongest. But as the devices move, especially when the relative position between the tag and the spike camera changes rapidly or unpredictably, the key point might move in the field of view. Therefore, we design a tracking algorithm to update the position of the key point in nearby pixels. Given the sufficient sampling rate of the spike camera, an acceptable key point can be found in the neighboring pixels of the current key point. If the key point can be tracked accurately and timely, we can further robustly demodulate the signal when the devices are mobile and the environment is changing.

During communication, we first manually select the estimated region(s) of the shutter(s) in the FoV of the camera, which can be further integrated with object detection algorithms [35]. The tracking algorithm then selects an initial key point by choosing the point with the best signal quality among the points with signal quality above a certain threshold *thres_init*. The demodulation algorithm then uses the initial key point to demodulate the first symbol. Afterward, when a point around the prior key point has a higher signal quality than the prior key point by a threshold *thres_track*, the point with the highest signal quality in the neighborhood will then become the new key point.

4 EVALUATION

4.1 Experimental Setup

As shown in Fig. 3a, our VLBC tag is divided into 8 control units, and different units modulate messages independently, *i.e.,* via spatial multiplexing. The power consumption of the tag is 0.8 mW during transmission, and a 6-watt LED enhances the brightness of the tag. The tag transmits random coded messages (distributed to the control units) with variant payload lengths and is captured by the



(a) Data rate. **(b) Packet length.** **(c) Mobility.** **(d) Communication range.**

Figure 4: Evaluation results.

spike camera. The spike camera has a spatial resolution of 250×400, F/1.4 lens, and 20000 Hz sampling rate. By default, we set $\delta = 64$, $thres = 5$, $\lambda = 0.9^{1/period}$ where $period$ is the number of images in one bit, $N = 2$, $M = 32$. $\alpha = \beta = 1$, $thres\_init = thres\_track = max\_quality/3$ where $max\_quality$ is the maximum possible signal quality, and the *log* kernel for the grayscale transformation. We mainly evaluate the performance of NeuromorphicVLC in terms of bit error rate, with respect to data rate (the sum of 8 control units), frame length, mobility, and communication range.

4.2 Experimental Results

Data Rate. The experimental setup is shown in Fig. 3b. We place the spike camera 50 cm from the VLBC tag, which sends 64-byte messages at frequencies from 0.8 Kbps to 6.4 Kbps (100 to 800 bps per control unit). We collect 2000 bytes of messages for each data rate and calculate the bit error rate. The results are shown in Fig. 4a. At 0.8 to 3.2 Kbps, our system works with negligible error in the setting. The performance of the system degrades rapidly when the data rate exceeds 4.8 Kbps. A thorough state transition of liquid crystal costs about 4 ms (250 Hz), whereas a higher data rate incurs insufficient state transition and smaller brightness difference between the states. Thanks to the adaptive quantization algorithm, our system can capture these smaller changes in brightness when the frequency is less than 600 Hz. The spatial resolution of a camera further enables 8 concurrent links to turboboost the overall throughput. This result indicates that the spike camera enhances the throughput of VLC because of its simultaneous high-frequency reception (KHz-level) of multiple concurrent links (pixels). The spike cameras demonstrate superiority over standard cameras (which suffer from low frame rates) and photodiodes (which do not have spatial-domain parallel reception and support up to 1 Kbps with one photodiode [23]).

Data rate	Bit error rate			Packet error rate		
	4.0 Kbps	4.8 Kbps	5.6 Kbps	4.0 Kbps	4.8 Kbps	5.6 Kbps
8 byte	0.06%	0.04%	1.41%	1.67%	3.57%	53.12%
16 byte	0.00%	0.07%	0.91%	0.00%	3.57%	55.00%
24 byte	0.04%	0.14%	0.42%	2.27%	8.33%	42.86%
32 byte	0.00%	0.04%	0.69%	0.00%	12.50%	55.56%

Table 2: Error rates.

Frame Length. We send a payload of varying lengths from 8 bytes to 32 bytes, with data rates ranging from 4.0 Kbps to 5.6 Kbps. The result is shown in Tab. 2. The detailed bit error rate distribution of 4.8 Kbps is shown in Fig. 4b. The bit error rate and packet error rate increase as the packet length increases. However, within the data rate of 4.8 Kbps, the bit error rate is still less than 10^{-3}, which is within the correction ability of error-correcting codes [36].

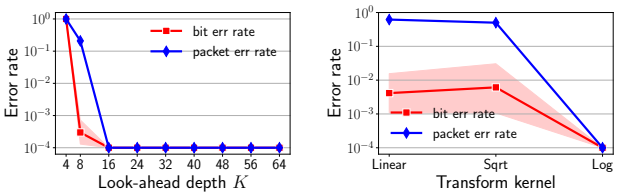
Mobility. We measure the performance of the system in different mobile situations. The experimental setup is shown in Fig. 3c. We place the camera on a cardboard box at a distance of about 1 meter from the tag. We tested six kinds of mobile cases: *static*, *far*, *moving*, *fast moving*, *shaking*, and *turning*. In the *static* and *far* cases, the camera is 1 m and 2 m away from the tag. In the *moving* and *faster* cases, we move the tag forward by 0.25 m and 0.5 m in 1 second[1]. We swing the camera left and right in the *turning* case, and pan the camera up and down in the *shaking* case. As shown in Fig. 4c, in the most common scenes of vehicle movement in reality (*i.e., moving* and *fast moving*), the performance of our system is similar to the *static* case. The performance gradually decreases as the motion becomes stronger and more complex. *Shaking* and *turning* have the greatest impact on performance. We attribute the phenomenon to abrupt changes in illumination patterns, which are to be solved with a more robust demodulation algorithm.

Communication Distance. We scale down the image captured from 64-byte payloads and 4.0-Kbps transmission, to simulate different real-world communication distances. To establish the relationship between the number of pixels for each control unit and the real-world distance, we use the parameters from a future spike camera that has the same pixel density as a commercial smartphone camera, with a FOV of 119°, an aspect ratio of 4:3, and a horizontal pixel count of 4000. Such an advanced spike camera is not available currently, but we believe this assumption is reasonable because spike cameras essentially share similar optics and circuit designs as standard CMOS vision sensors of existing smartphones. We calculate the equivalent distance of a 1-meter-wide square street sign as the VLBC tag in the experiment. The results are shown in Fig. 4d. In the stationary case, the bit error rate is hardly affected by a distance of up to 800 meters. In contrast, the performance is still robust in 100 meters in the moving case but degrades rapidly because of frequent failure of key point tracking. Despite the low resolution (potentially from a distant tag) of the image, the message can still be demodulated as long as the number of effective pixels of the LCD tag is sufficient for key point tracking.

4.3 Microbenchmarks

We evaluate how the parameters of our image reconstruction algorithm influence the error rate performance of demodulation with two microbenchmarks.

[1]We would like to note that the relative speed in the FoV of the camera is the most important factor in the performance. The tested speeds can be equivalent to 90 and 180 km/h when the communication distance is scaled to 100 m, although the in-lab speeds might seem slow because the distance is only about 1 m.



(a) Look-ahead depth K. **(b) Kernels.**

Figure 5: Microbenchmarks.

Look-ahead Depth K. The hyperparameter K in §3.2 indicates the depth of the spike stream used for grayscale image reconstruction. It affects the minimum brightness of the reconstructed grayscale image in our reconstruction algorithm. We use different K settings to reconstruct and further demodulate the experiment data sent at 4.0 Kbps, and show the results in Fig. 5a. In our setting, we observe that a spike is generated every 10 ~ 15 samples in the spike stream when the tag is dark (*i.e.,* the liquid crystal shutter blocks light to pass), and the interval decreases to 2 ~ 5 when the tag becomes bright. Therefore, when K is set larger than 16, the BER performance is not influenced by our reconstruction algorithm that omits the dark areas. In practical applications, K (and the camera aperture) should be adjusted properly to control the amount of incident light for effective and efficient image reconstruction.

Grayscale Transformation Kernel. As illustrated in §3.2, the grayscale transform kernel affects the contrast of the image. The tags are generally brighter than the nearby environment thanks to retroreflection. Therefore, a kernel with a larger degree of nonlinearity would increase the contrast in the bright parts of the image, which is typically the area of interest for communication thanks to retroreflection. As a result, the *log* kernel performs best in our experiments as shown in Fig. 5b.

5 CONCLUSION AND OPPORTUNITIES

In this paper, we proposed and prototyped NeuromorphicVLC that leverages a neuromorphic camera to improve the performance of vehicular VLC. We have shown that the neuromorphic camera, as a novel class of optical receivers, can make the best of the two conventional VLC receivers, *i.e.,* photodiodes and standard cameras. Although NeuromorphicVLC is primarily designed for V2X networks, the introduction of neuromorphic cameras would also benefit a wider range of (indoor stationary) VLC systems by enabling massive concurrent communication links, which chimes with the massive MIMO technology in the RF domain [37]. In-lab experimental results have preliminarily demonstrated the feasibility of NeuromorphicVLC. Future research opportunities may include:

Communication Distance. An adequate distance of communication is critical to allow the vehicles to have enough space and time to take action. In addition to the number of available pixels evaluated in this paper, the limited resolution of optical imaging from atmospheric disturbance, inaccurate focusing, and imperfect lenses might also be bottlenecks of communication distance. A novel scanning light-field imaging sensor [38] can potentially alleviate this issue. Besides, a vehicle may generate high-frequency vibrations that would incur motion blurring of the image and can be potentially handled with advanced deblurring algorithms [39].

From Communication to Networking. In a V2X network, multiple tags may simultaneously appear in the FoV of one vehicle. An effective multiple access scheme would be important to the utility of the channel. Thanks to the imaging capability of cameras, the tags can be distinguished from each other in the spatial domain, so a straightforward multiple access scheme would work well, at a cost of extra complexity of the demodulation algorithm. Various active light sources (*e.g.,* LCD screens and LEDs) may also function as VLC transmitters and appear simultaneously in the network. Such active transmitters may benefit reliability because they are less affected by ambient illumination. However, they have high power consumption (require external power source) and thus are less applicable to the existing traffic sign infrastructure, especially in rural areas. A well-designed network should coordinate both active and passive (*e.g.,* VLBC tags in this paper) transmitters to take full advantage of the two types of devices.

Integrated Sensing and Communication. State-of-the-art autonomous vehicles typically use cameras to perform computer vision perception. If VLC is to be received by a conventional camera in a real-world application, both the sensing and the communication performance would be degraded due to the rolling-shutter effect, and can only be alleviated with a dedicated machine learning algorithm [40] to our knowledge. NeuromorphicVLC, without the impact of the rolling-shutter effect, presents a more fundamental step towards integrated sensing and communication. With the high temporal resolution (oversampling), neuromorphic cameras have the ability to separate the mutual interference between communication and sensing (perception) towards an integrated design.

Real-world Deployment. Real-world deployment may pose more challenges to NeuromorphicVLC and any other vehicular VLC systems with similar settings. There are numerous factors that may influence the performance of vehicular VLC, such as relative orientations, mobile (speed, turning, shaking, phase misalignment, *etc.*) scenarios, adverse weather/lighting conditions and sudden changes, background luminous and illuminated objects (which may also be moving), and random occlusions. We believe more experiences in the deployment of vehicular VLC systems are helpful to discover more practical challenges and raise corresponding solutions. We also believe a large-scale real-world benchmark would help the community to evaluate different systems with substantive quantitative backing. However, such a benchmark is still missing in today's literature to our knowledge, and we believe it would greatly help the community to explore this domain.

ACKNOWLEDGMENTS

We are grateful to our shepherd, Dr. Longfei Shangguan, and the anonymous HotMobile reviewers for their constructive critique and valuable comments, all of which have greatly helped us improve this paper. This work is supported in part by the National Key Research and Development Plan, China (Grant No. 2020YFB1710900) and the National Natural Science Foundation of China (Grant No. 62022005, 62272010, 62061146001, 62088102 and 62136001). Chenren Xu is the corresponding author.

REFERENCES

[1] Wireless access in vehicular environments. *IEEE Std 802.11p*, 2010.

[2] Shanzhi Chen, Jinling Hu, Yan Shi, Ying Peng, Jiayi Fang, Rui Zhao, and Li Zhao. Vehicle-to-everything (V2X) services supported by LTE-based systems and 5G. *IEEE Communications Standards Magazine*, 2017.

[3] Mohsen Kavehrad. Optical wireless applications: A solution to ease the wireless airwaves spectrum crunch. In *SPIE Broadband Access Communication Technologies Vii*, 2013.

[4] Behnaam Aazhang et al. *Key drivers and research challenges for 6G ubiquitous wireless intelligence (white paper)*. 2019.

[5] Wei Jiang and Fa-Long Luo. Optical and visible light wireless communications in 6G. 2023.

[6] Luiz Eduardo Mendes Matheus, Alex Borges Vieira, Luiz FM Vieira, Marcos AM Vieira, and Omprakash Gnawali. Visible light communication: concepts, applications and challenges. *IEEE Communications Surveys & Tutorials*, 2019.

[7] Airbus, Latécoère bring LiFi IFE to the business jet market. https://paxex.aero/airbus-latecoere-lifi-business-jet/, 2021.

[8] IEEE standard for local and metropolitan area networks–part 15.7: Short-range optical wireless communications. *IEEE Std 802.15.7*, 2019.

[9] High-speed indoor visible light communication transceiver - system architecture, physical layer and data link layer specification. *ITU-T G.9991*, 2019.

[10] Agon Memedi and Falko Dressler. Vehicular visible light communications: A survey. *IEEE Communications Surveys & Tutorials*, 2020.

[11] Alin-Mihai Căilean and Mihai Dimian. Current challenges for visible light communications usage in vehicle applications: A survey. *IEEE Communications Surveys & Tutorials*, 2017.

[12] Pengfei Luo, Min Zhang, Zabih Ghassemlooy, et al. Experimental demonstration of RGB LED-based optical camera communications. *IEEE Photonics Journal*, 2015.

[13] Peng Ji, Hsin-Mu Tsai, Chao Wang, and Fuqiang Liu. Vehicular visible light communications with LED taillight and rolling shutter camera. In *IEEE Vehicular Technology Conference*, 2014.

[14] Takaya Yamazato, Isamu Takai, Hiraku Okada, et al. Image-sensor-based visible light communication for automotive applications. *IEEE Communications Magazine*, 2014.

[15] Wen-Hsuan Shen and Hsin-Mu Tsai. RayTrack: enabling interference-free outdoor mobile vlc with dynamic field-of-view. In *ACM MobiSys*, 2021.

[16] Shinya Itoh, Isamu Takai, Md Shakowat Zaman Sarker, Moeta Hamai, Keita Yasutomi, Michinori Andoh, and Shoji Kawahito. A CMOS image sensor for 10Mb/s 70m-range led-based spatial optical communication. In *IEEE ISSCC*, 2010.

[17] Fuyou Liao, Feichi Zhou, and Yang Chai. Neuromorphic vision sensors: Principle, progress and perspectives. *Journal of Semiconductors*, 2021.

[18] Guillermo Gallego, Tobi Delbrück, Garrick Orchard, et al. Event-based vision: A survey. *IEEE TPAMI*, 2020.

[19] Wen-Hsuan Shen, Po-Wen Chen, and Hsin-Mu Tsai. Vehicular visible light communication with dynamic vision sensor: A preliminary study. In *IEEE VNC*, 2018.

[20] Zhengqiang Tang, Takaya Yamazato, and Shintaro Arai. A preliminary investigation for event camera-based visible light communication using the propeller-type rotary LED transmitter. In *IEEE ICC Workshops*, 2022.

[21] Tiejun Huang, Yajing Zheng, Zhaofei Yu, et al. 1000x faster camera and machine vision with ordinary devices. *Engineering*, 2022.

[22] Lin Zhu, Siwei Dong, Jianing Li, Tiejun Huang, and Yonghong Tian. Ultra-high temporal resolution visual reconstruction from a fovea-like spike camera via spiking neuron model. *IEEE TPAMI*, 2022.

[23] Xieyang Xu, Yang Shen, Junrui Yang, Chenren Xu, Guobin Shen, Guojun Chen, and Yunzhe Ni. PassiveVLC: Enabling practical visible light backscatter communication for battery-free IoT applications. In *ACM MobiCom*, 2017.

[24] Purui Wang, Lilei Feng, Guojun Chen, et al. Renovating road signs for infrastructure-to-vehicle networking: A visible light backscatter communication and networking approach. In *ACM MobiCom*, 2020.

[25] Ashwin Ashok, Marco Gruteser, Narayan Mandayam, Jayant Silva, Michael Varga, and Kristin Dana. Challenge: Mobile optical networks through visual MIMO. In *ACM MobiCom*, 2010.

[26] Cen B Liu, Bahareh Sadeghi, and Edward W Knightly. Enabling vehicular visible light communication (V^2LC) networks. In *ACM VANET*, 2011.

[27] Alin-Mihai Căilean, Barthélemy Cagneau, Luc Chassagne, Mihai Dimian, and Valentin Popa. Novel receiver sensor for visible light communications in automotive applications. *IEEE Sensors Journal*, 2015.

[28] Liwen Hu, Rui Zhao, Ziluo Ding, Lei Ma, Boxin Shi, Ruiqin Xiong, and Tiejun Huang. Optical flow estimation for spiking camera. In *IEEE/CVF CVPR*, 2022.

[29] Jiyuan Zhang, Lulu Tang, Zhaofei Yu, Jiwen Lu, and Tiejun Huang. Spike transformer: Monocular depth estimation for spiking camera. In *ECCV*, 2022.

[30] Jing Zhao, Jiyu Xie, Ruiqin Xiong, Jian Zhang, Zhaofei Yu, and Tiejun Huang. Super resolve dynamic scene from continuous spike streams. In *IEEE/CVF ICCV*, 2021.

[31] Yue Wu, Purui Wang, Kenuo Xu, Lilei Feng, and Chenren Xu. Turboboosting visible light backscatter communication. In *ACM SIGCOMM*, 2020.

[32] Kenuo Xu, Chen Gong, Bo Liang, Yue Wu, Boya Di, Lingyang Song, and Chenren Xu. Low-latency visible light backscatter networking with RetroMUMIMO. In *ACM SenSys*, 2022.

[33] Wei Li, Tuochao Chen, Zhe Ou, Xin Wen, Zichen Xu, and Chenren Xu. RetroFlex: enabling intuitive human–robot collaboration with flexible retroreflective tags. *CCF Transactions on Pervasive Computing and Interaction*, 2022.

[34] Lin Zhu, Siwei Dong, Tiejun Huang, and Yonghong Tian. A retina-inspired sampling method for visual texture reconstruction. In *IEEE ICME*, 2019.

[35] Etienne Perot, Pierre de Tournemire, Davide Nitti, Jonathan Masci, and Amos Sironi. Learning to detect objects with a 1 megapixel event camera. In *NeurIPS*, 2020.

[36] Abdelhamid Nafaa, Tarik Taleb, and Liam Murphy. Forward error correction strategies for media streaming over wireless networks. *IEEE Communications Magazine*, 2008.

[37] Clayton Shepard, Josh Blum, Ryan E Guerra, Rahman Doost-Mohammady, and Lin Zhong. Design and implementation of scalable massive-MIMO networks. In *ACM MobiSys Workshops*, 2020.

[38] Jiamin Wu, Yuduo Guo, Chao Deng, Anke Zhang, Hui Qiao, Zhi Lu, Jiachen Xie, Lu Fang, and Qionghai Dai. An integrated imaging sensor for aberration-corrected 3D photography. *Nature*, 2022.

[39] Zhe Jiang, Yu Zhang, Dongqing Zou, Jimmy Ren, Jiancheng Lv, and Yebin Liu. Learning event-based motion deblurring. In *IEEE/CVF CVPR*, 2020.

[40] Ziwei Liu, Tianyue Zheng, Chao Hu, Yanbing Yang, Yimao Sun, Yi Zhang, Zhe Chen, Liangyin Chen, and Jun Luo. CORE-Lens: Simultaneous communication and object recognition with disentangled-gan cameras. In *ACM MobiCom*, 2022.