

Image Sensor Based V2V Optical Wireless Communication System

Chiaki AOYAMA*

ABSTRACT

In order for multiple autonomous vehicles to interact, they will need to identify one another according to their relative position and ID. With this objective, a camera-based system for vehicle-to-vehicle communication was developed to detect vehicle IDs and relative position by imaging optical signals loaded with an ID and blinking at high speed. Digital information is sent with modulated light that changes in phase, and phase changes are detected by lock-in pixel camera to receive the signals. The system also detects distance by shifting the relative sampling timing. The implementation can receive 640 bps signals simultaneously on seven channels and detect distances up to 600 m with an uncertainty of ± 5.6 m.

1. Introduction

With the advances in IT, various parties are currently working on implementation of autonomous vehicles. In order for them to move autonomously without colliding, such vehicles are equipped with multiple distance sensors, including stereo cameras, millimeter-wave radar, and laser rangefinders, to detect their surrounding environment.

In addition to the progress being made on standalone vehicles, vehicle-to-vehicle (V2V) communication is also evolving toward making traffic flow more smoothly and safely, allowing vehicles to alternate when merging into traffic, and to accelerate and decelerate without delay when following another vehicle. In Japan, the Ministry of Internal Affairs and Communications (MIC) is heading up research on systems utilizing V2V communication and roadside-to-vehicle (RTV) communication⁽¹⁾. Accordingly, systems using wireless communications and position information from the Global Navigation Satellite System (GNSS) are being researched (see Fig. 1).

These cooperative systems require communications that can sense relative positions accompanying individual identification by vehicle IDs.

Thus, stabilizing detection for relative positions during communications is an issue. The cooperative systems are built to constantly receive vehicle information for multiple vehicles within an approximately 300-m radius at a frequency of 10 Hz^{(2), (3)}. This vehicle information includes details such as vehicle position, speed, level of acceleration or deceleration, and traveling direction. In order to select vehicle information for a specific position among multiple vehicles, the system determines the information for a vehicle

in a specific position from the positions for all received vehicle information by comparing it with the system's own position. As the system uses GNSS to detect position, positions will be detected at reduced accuracy or not at all if the vehicle cannot be observed by multiple satellites due to the way GNSS detects positions. When downtown or in tunnels, this makes coordinated driving less feasible.

Distance sensors can be used to detect relative position, but have issues identifying individual vehicles. If the sensors could identify individual vehicles, they could obtain pre-registered vehicle addresses and communicate based on these addresses. However, distance sensors are for measuring relative positions up to approximately 100 m, and even if they could separate the detected objects based on their shape and movement by making detection points finer, individual vehicle identification is not possible without the ability to estimate license plates from the vehicle shape.

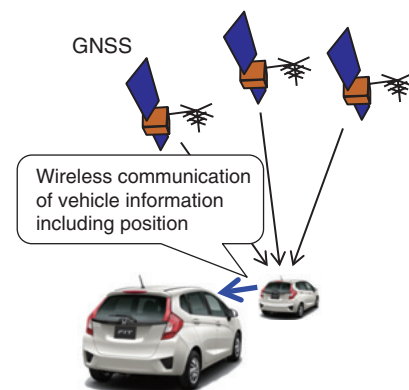


Fig. 1 V2V communication

* Fundamental Technology Research Center

Thus, communication with a specific vehicle is not possible through any combination of distance sensors and mobile communication or near-field communication systems.

If a system were capable of both detecting relative position and identifying individuals, it could communicate with specific vehicles through next-generation technology like 5G and other various means of wireless communication. For example, if there were a system to read vehicle license plates and search addresses through the license plate numbers, it could communicate with any vehicle with readable license plates. By reading the license plate numbers from images, the system could also detect relative position. Trying to read numbers from long distances, however, requires a number reader with a telephoto camera as part of a complex system with anti-vibration and swivel mechanisms. Also, sensing vehicles in the neighboring lane would require license plate numbers mounted on the vehicle sides.

This paper therefore reports the evaluated performance of a system constructed using a camera to receive signals sent from transmitters with LED light sources (see Fig. 2), allowing it to simultaneously identify individual vehicles and detect their relative position.

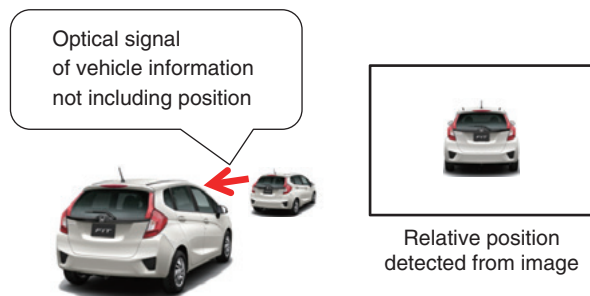


Fig. 2 Optical wireless communication camera

2. Prior Art and Issues

The camera-based methods of individual identification and position detection are each explained below.

In terms of individual identification, research has advanced on a method of camera-based optical wireless communication that changes the light source to decode lower resolution signals⁽⁴⁾⁻⁽⁸⁾. The system will be required to communicate simultaneously with the preceding vehicle and vehicles in the next lanes, and needs to be able to communicate with at least three vehicles simultaneously. It also needs to communicate with multiple light sources when detecting its own position from the light source position.

In terms of relative position detection, as stereoscopic views and other distance sensors can be used for short distances within approximately 100 m, the system uses a method that can detect relative position at longer distances. In terms of detection distances, the system requires 400 m for communicating with oncoming traffic when overtaking low-speed vehicles. As direction can be detected from

image positions if a camera is used, relative position can be detected if the distance is also detected simultaneously.

2.1. Optical Communication Method

Optical communication methods for receiving optical signals by camera (a two-dimensional light-receiving element) are roughly divided by camera types: general-purpose cameras, compound-eye cameras, direct readout cameras, and lock-in pixel cameras.

With general-purpose camera communication, the camera continuously captures optical signals, tracking the changes in light intensity sequentially by time and demodulating the signals⁽⁴⁾. As the cameras can normally capture images at 30 to 60 FPS, the time resolution for signals is roughly 16 to 33 ms. Thus, assuming three images are needed to decode a 1-bit signal, the signal rate will be 5 to 11 bits per second.

If the general-purpose color camera continuously captures optical signals, changes in light intensity can be tracked sequentially by time, and the signals then demodulated⁽⁵⁾. As the human eye is also sensitive to color changes, physical and psychological impacts will be the issue if multiple light sources are put in place.

There is a method of restoring signals from the striped scan line patterns generated when imaging a surface-emitting light source of varying intensity using a general-purpose rolling-shutter camera, as often seen in smartphones⁽⁶⁾. The light-emitting source needs to be of a certain length and perpendicular to the scan lines. This places size limitations on the light source, which is an issue for communications at a distance.

With compound-eye camera communication, elements capable of receiving optical signals are arranged two-dimensionally⁽⁷⁾. In this camera, a lens is placed in front of a two-dimensionally arranged light-receiving section of an infrared remote controller. This form of optical communication would be high speed as it can be used without alteration, but it is difficult to increase resolution.

With direct readout camera communication, the current generated by light incident to pixels is continuously retrieved at an external source, allowing conventional optical signal decoding technology to be applied without alteration⁽⁸⁾. This method requires a special imaging element. As the electrons generated by photoelectric conversion is continuously being read out as a current, sensitivity is low, and it is not easy to communicate at distance with a wide-field camera. Furthermore, communicating with multiple light sources is difficult.

With lock-in pixel camera communication, the camera detects the amplitude and phase of modulated light by synchronizing with the modulated light frequency, then storing the electrons generated by a photoelectric converter in different charge accumulators⁽⁹⁾. In this research, this was the method adopted to decode a phase-altered signal. This method can detect modulated light at frequencies exceeding camera frame rates. With it, frequencies can be set high and only the phase has to be changed, so the human

eye will not perceive any blinking. In addition, as it uses modulated light, candidate signal sources can easily be narrowed down to within a single frame. Also, as electric charge is integrated before detection, sensitivity is higher than with the direct readout method. While the transmission speed is slower than with direct readout, speed can be increased system-wide through use in conjunction with high-speed wireless technology, such as 5G, thus giving it the sensitivity to be usable at long distances. This explains our selection of this method.

2.2. Distance Detection Method

Camera-based methods for detecting distance are broadly classified into monoscopic vision, stereo vision, and Time of Flight (TOF).

In the monoscopic vision method, distance is detected from the location of features on the image of a subject whose position has already been determined⁽¹⁰⁾. Accuracy is good if multiple features are detected with sufficient distance between them.

Stereo vision detects distance based on the location of a single subject in images taken from multiple points of view⁽¹¹⁾. A wider camera separation or higher resolution is required to maintain measurement accuracy at longer distances.

TOF detects distance based on the time light requires to move through space. Here, the difference between when the light departs and when it enters the camera is multiplied by the speed of light to convert to distance. TOF was selected for use in this research as subject size is not a factor and distance can be detected up to a few hundred meters.

TOF is further subdivided into lock-in pixel camera TOF and draining-only modulation camera TOF.

The lock-in pixel camera method uses a pixelated camera with capacitor sections to store multiple electric charges and a photoelectric conversion section to convert light into electric charge. Here, the generated charges are distributed to the capacitor sections at differed timings. Distance is then detected by calculating the time required for the light

to move through space from the ratios of the number of charges stored in the multiple capacitor sections⁽¹²⁾. In order to detect the correct distances, the characteristics of each capacitor section have to be calibrated according to the reflected light intensity, ambient light intensity, and temperature. Calibrating for three parameters is not easy.

A draining-only modulation camera⁽¹³⁾ detects distance by repeatedly imaging for exposure periods of 100 ns or less. The photoelectrons can be discarded at arbitrary timing using the drain function. These cameras can detect the distance similar to the lock-in pixel camera by changing timing with multiple pixels. Compared to the lock-in pixel camera, sensitivity at the same area is reduced due to the discarded loads. By changing the timing in chronological order and imaging multiple times, this method obtains the same results as the lock-in pixel camera without changing the timing between pixels. As light source intensity is higher imaging direct light than it is in reflected light, allowing detection at lower sensitivity, the draining-only system was used.

In summation, the system adopted for this research uses three capacitor sections for phase detection, leaving the remaining capacitor section available to detect distance.

3. Fundamentals of the System

This system consists of a light projector to project modulated light carrying a signal, camera-based light receivers to decode the signals from the modulated light and detect the TOF pulses for distance detection, a modulator to send the signals to the light receivers, and a demodulator and oscillator to decode the signals (see Fig. 3). With frequency stabilized through GNSS, the oscillator enhances distance detection performance and, if times are synchronized, decreases the delay times in detection. When GNSS data is not available, the system is still capable of detection, however with longer detection times.

3.1. Pixel Functions

The pixel circuit for a single pixel is constructed as

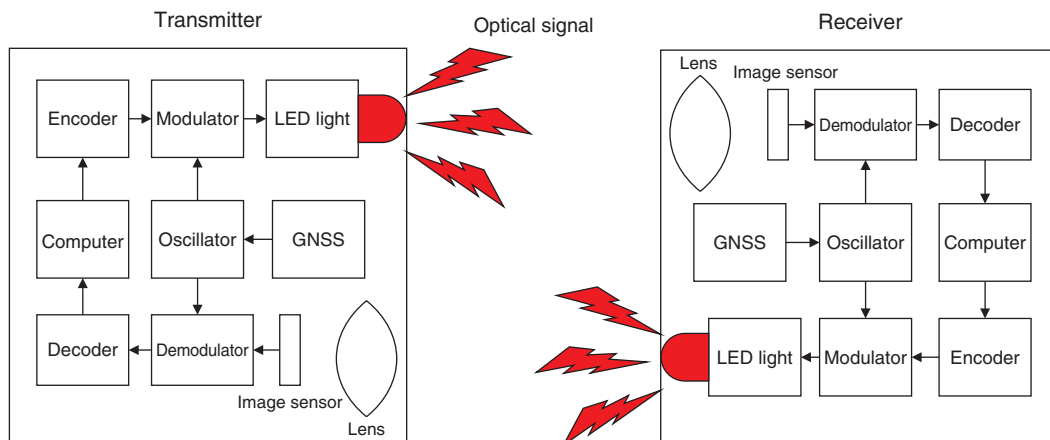


Fig. 3 System configuration

shown in Fig. 4. The photoelectric conversion section converts light into electric charge. The ends of the transfer gates are connected to the capacitor sections, and the generated charges are stored (integrated) each time a transfer gate opens. The end of the drain gate is connected to the power source and purges the generated charges. Charges stored in the capacitor sections are converted to current by the readout circuit and read out of the pixel.

Three out of the four capacitor sections are used for communication, with the remaining one used for detecting distance. Transfer gates TX1, TX2, and TX3 are sequentially opened at staggered timings, storing charges a total of eight times in their respective capacitor sections. The transfer gate cycle for gates TX1, TX2, and TX3 includes a brief period when gate TX4 is opened, during which the other transfer gates are closed (see Fig. 5).

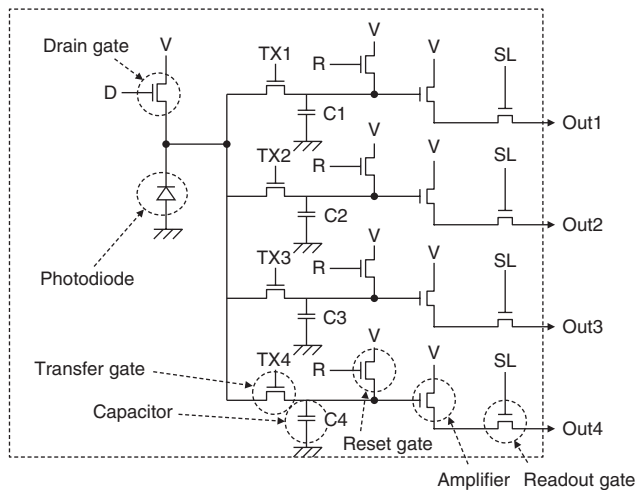


Fig. 4 Pixel circuit

3.2. Transmission Waveform

The signal pulse of 36- μ s cycles (27.8 kHz) as used for communication is taken as the basic waveform SP for transmission (see Fig. 6). This pulse reiterates a consecutive 24 times, then rests for eight empty cycles. This 32-cycle period is taken as one bit period, called a signal block (SB). When the signal pulse phase is changed from the previous block, the change is expressed in binary with a 0 or 1. When 0, the phase is advanced by $2/3\pi$ (120°); when 1, it is retarded by $2/3\pi$ (see Fig. 8). Thus, as long as the phase stays within the range of $\pm 60^\circ$, noise will not result in an error. The start of the signal sequence is made leader by continuing at the same phase. Subsequently, sections with changing phases are expressed in binary, ending with a period with no signals. These pulse sequences of 84 pulses with two leaders and two periods of no signals comprise a single signal sequence. In order to fill the sequence cycle out into the 100 ms⁽²⁾ required for V2V communications, an additional adjustment period is inserted after the periods of no signals. Thus, the system can receive 10 IDs per second.

The 36- μ s basic waveform includes the 200-ns pulse signal sequence used for detecting distances. The timing for the start of the pulse sequences shifts by 25 ns for each signal block, bringing it back to the original position every 80 iterations.

3.3. Optical Signal Detection Method

The phase and amplitude for the pulsed light are derived from the values sampled for one cycle of modulated light in the three capacitor sections (see Fig. 7). As levels are insufficient after just one sampling, however, eight samples are taken for each cycle to calculate the phase and amplitude as the integrated value of the signal of the same timing.

Phase changes are detected and decoded to binary for the binary signals (see Fig. 8). Phase is assigned in $2/3\pi$ increments to limit error across three samples.

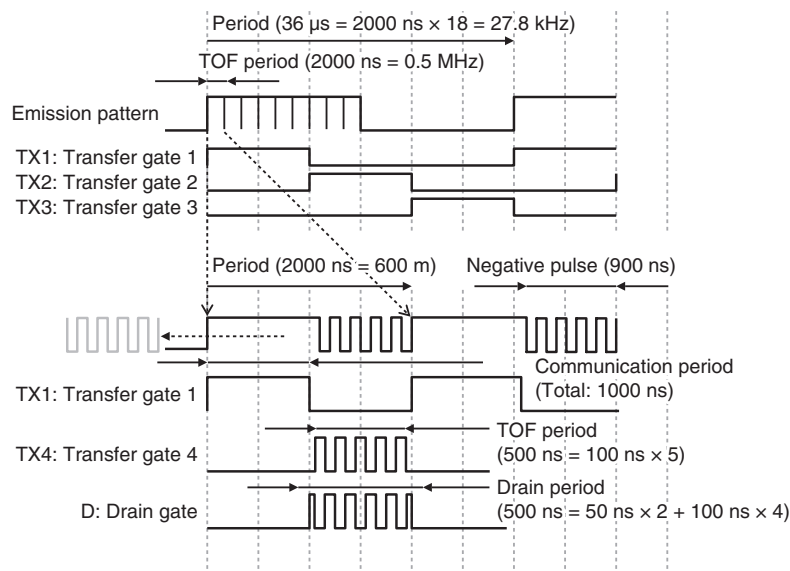


Fig. 5 Transfer gate timing

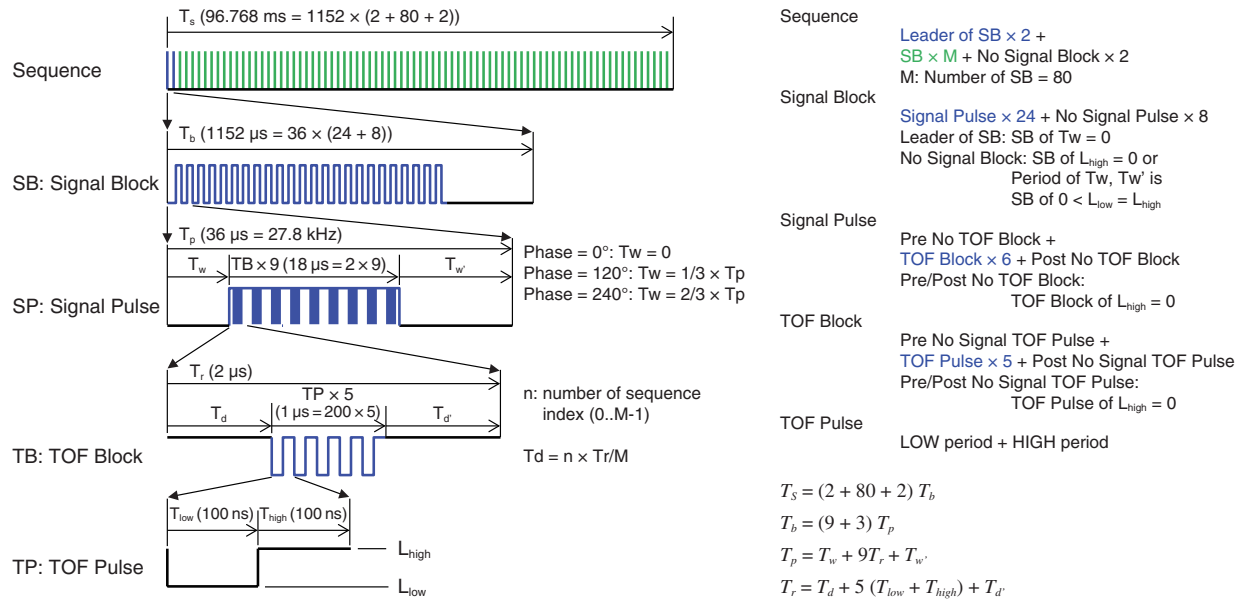


Fig. 6 Wave configuration

Detecting phase θ of three integrated values obtained by shifting integration period

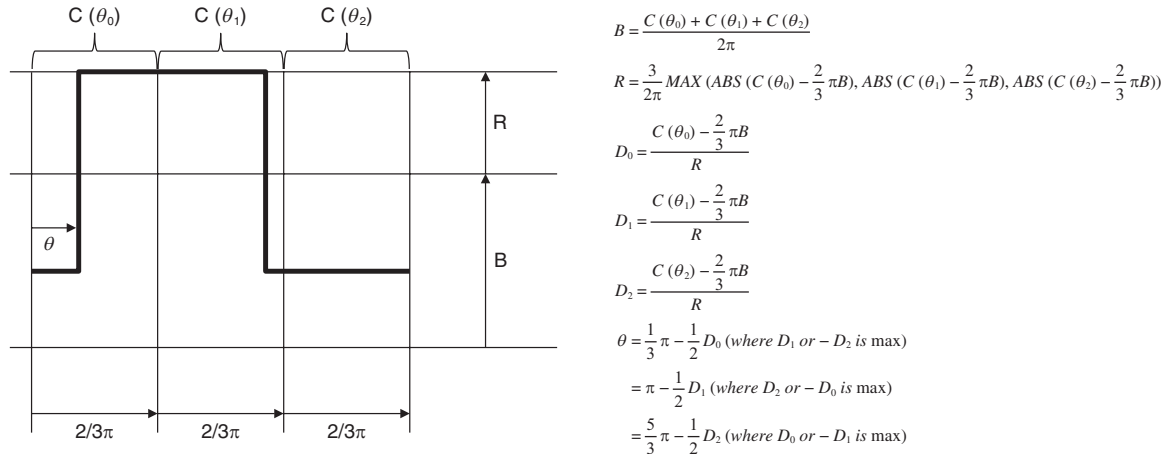


Fig. 7 Phase calculation

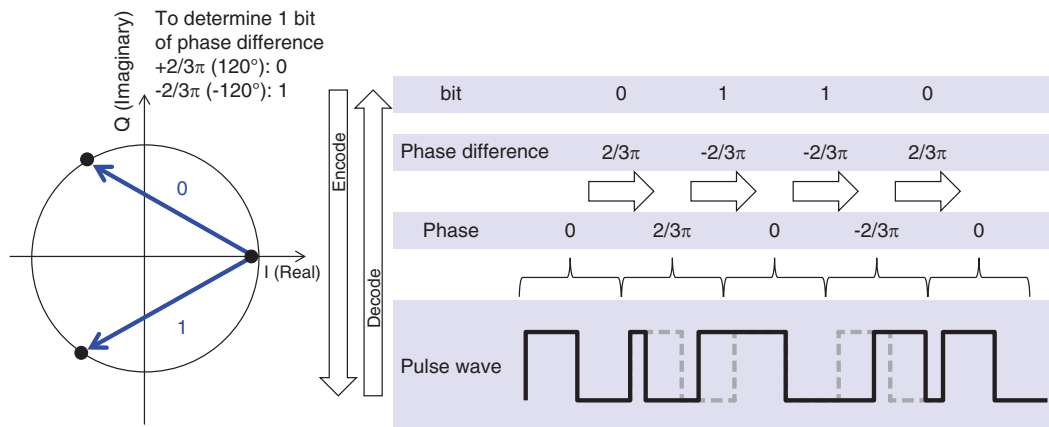


Fig. 8 Encoding and decoding process

3.4. Distance Detection Method

In order to permit measurement of longer distances, the distance detection method involves systems communicating with each other via light-emitting sources, similar to secondary surveillance radar⁽¹⁴⁾ as used in aircraft (see Fig. 9). When detecting distance with reflected light, light intensity decreases in inverse proportion to distance to the fourth power when the size of the light source as captured on the image is one pixel or smaller. With direct light, however, intensity decreases in inverse proportion to the distance squared, giving it an advantage

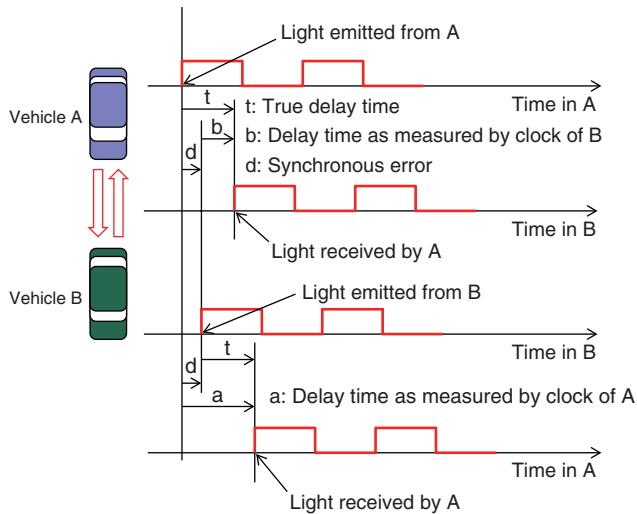


Fig. 9 Distance calculation

at longer distances. Unlike secondary surveillance radar, signals cannot be sent back after a set delay period from when the signal was received as the signals are sampled discretely. Therefore, each system sends the other signals continuously and independently, communicating to one another the delay times as they are detected in order to cancel out synchronization error and calculate the true delay time. Taking b as the delay time measured by the clock of Vehicle B when Vehicle B receives the pulsed light emitted by Vehicle A, a as the delay time when Vehicle A receives the pulsed light emitted by Vehicle B, and d as the synchronization error between the clocks of the two vehicles, the true delay time t can be solved as:

$$t = b + d$$

$$t = a - d$$

Canceling out the unknown d , t can thus be calculated as:

$$t = (a + b)/2$$

In measuring delay time for the pulsed light, the relative sample timings will gradually change, and delay time is calculated from the degree of coincidence. Gate output is at its lowest when the five gate timings coincide with five negative pulses, then increases with each shift in timing. Thus, the system finds the timing with the lowest output and sets it as the delay time (see Fig. 10). With single pulses, a short pulse increases accuracy in estimating peak position but reduces detection sensitivity, so pulses are sent in repeated bursts in order to increase sensitivity and detection accuracy. In addition, negative pulses are united with the

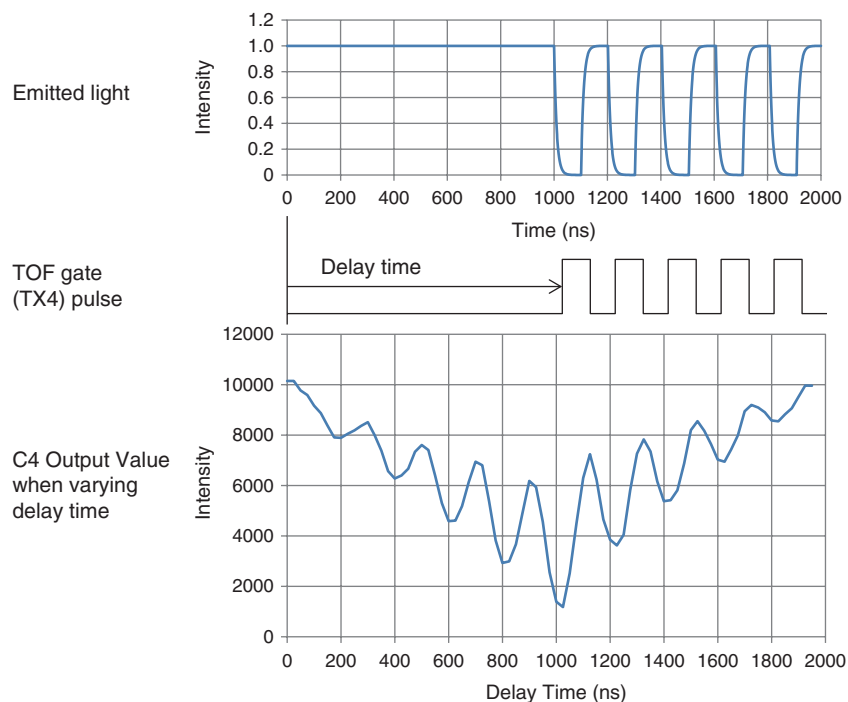


Fig. 10 Delay time detection

communication pulses to allow distance to be detected during communication without separating communication and distance measurements.

4. System Implementation

4.1. System Configuration

Each side of the system will consist of a communication distance camera, lens, projector, arbitrary waveform generator, GNSS oscillator, and calculator (see Fig. 11).

The GNSS oscillator outputs a 10-Hz pulse signal and 1-PPS pulse signal at synchronized times.

The arbitrary waveform generator outputs a clock signal used by a camera synchronized to 10 MHz and a communication distance waveform signal projected by the projector.

The communication distance camera consists of an imaging element, control board, and interface board. The imaging element sends image signals to a computer via the control board and interface board.

The projector emits LEDs as a communication distance optical signal, driven by light pulse signals input from the arbitrary waveform generator.

4.2. Pixel Structure

The same imaging element is used as in the differential camera⁽¹⁵⁾. It uses square pixels with 12- μm sides for an effective image size of 640 pixels horizontally and 480 vertically.

4.3. Region of Interest (ROI)

Image readout is broken into small regions to speed up the process. Up to eight small regions can be specified.

Regions can be subdivided into any number of detection regions and tracking regions. For this study, detection regions were set to a size of 64×24 pixels, scanning for regions of modulated light from top left to bottom right. When a region of modulated light is detected, the light is decoded into seven tracking regions of 8×8 pixels each. As the detection regions scan for modulated light at 27.8 kHz, no continuously bright sunlight or low-frequency blinking lights will qualify as candidate light sources.

4.4. Projector

For the projector, 10 near-infrared LEDs outputting approximately 1 W each were arranged in series across spaces of 940 nm and 860 nm. The half-value angle is approximately $\pm 35^\circ$, with no diffusion plate used. V2V communication is expected to use a 940-nm array, and RTV communication is expected to use an 860-nm array. As light would be visible in red at 860 nm and visible red light is prohibited by law, communication with visible light at 860 nm would be unsuitable. Meanwhile, 940-nm light results in a less sensitive imaging element, but is not visible. Thus, a 940-nm wavelength was selected for the on-board LED.

5. System Performance

5.1. Communication Performance

The system was operated with a 64-bit data component and 12-bit error correction with 4 bits of additional information. One sequence was 80-bit with no error correction. As 10 sequences per second can be received simultaneously across seven channels, the channels were 640 bps each for a total band of 4,480 bps. Only the ID was

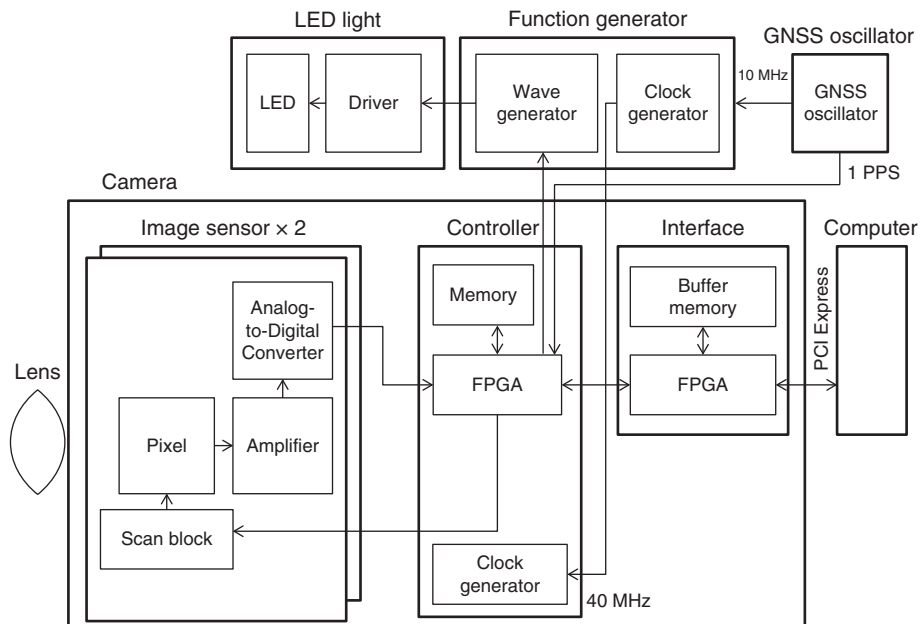


Fig. 11 Block diagram

listed in the data component. Depending on application, the data component will be loaded with a number indicating the data transmission order and vehicle speed information.

To measure communication performance, timing for a pulse generated by the arbitrary waveform generator was shifted over one cycle with the camera and projector set facing each other. Performance was then evaluated by shifting the detected pixel position.

Standard uncertainty for phase detection values was confirmed to be well within the $\pm 60^\circ$ range of binary detection at 5.4° or less within 1/8 intensity from the reference intensity.

A lens with a 56.5° horizontal angle of view and infrared transmission filter results in a maximum communication distance of 118 m outside in sunlight. Light at wavelengths of 800 nm and above can pass through the infrared transmission filter. Maximum communication distance is estimated to be expandable to 300 m or more just by narrowing the upper and lower irradiation angles with a lens or reflective mirror.

5.2. Distance Detection Performance

As with communication performance, distance detection performance was evaluated by changing the delay time and converting to distance.

Standard uncertainty for distance detection was 5.6 m or less within 1/8 intensity from the reference intensity. Distance was detectable within a range of 600 m. Variation due to temperature change is within ± 0.3 m at a temperature range of 5°C to 35°C . Calibration is unnecessary within this temperature range and likely even outside this range.

5.3. Tracking Performance

To evaluate performance in tracking a moving light source, the camera was loaded onto a rotating stage to evaluate performance in tracking a changing angular velocity.

When decoding signals in a small region with ROI, the light source was trackable at moving speeds of 480 pixels/s

and lower. For example, using a lens with a 60° horizontal angle of view, the camera could track a light source moving at an angular velocity of 68° per second. Given that the angular velocity of a vehicle traveling at a relative speed of 50 kph 1 second before passing is 23° per second, 68° is sufficient.

6. Sample Applications of the Communication Camera

As it will likely take some time for intercommunicating V2V and RTV communication systems to attain prevalence, an assumed application is automatic parking at homes, simply installing optical beacons locally as a positioning method.

Using optical beacons at known locations installed at seven points on interior walls, the system was tested for ability to calculate its own position (see Fig. 12). Using the optical beacons, the system was usable in both day and night settings.

Mounting the camera with a fisheye lens⁽¹⁶⁾ calibrated for a 180° horizontal angle of view, the camera detected position and posture of two or more light sources by detecting distance to the light sources from an elevated position in view of the light sources. As the camera was faced horizontally in anticipation of being mounted on vehicles, it detects position with just two or three light sources visible ahead. If the cameras are restricted to in-cabin mounting, the number of light sources can be reduced by pointing the camera toward the roof.

Standard uncertainty for position was 0.10 m, and that for angle was 0.64° . The camera requires 200 ms to determine position, at which point it can detect position at 250 Hz as long as it can track the light source. In terms of accuracy, the camera detects position as accurate as with existing methods⁽¹⁷⁾, while angle accuracy is an order of magnitude higher, and detection time was 1/5 that of existing methods. This demonstrates the method to be superior to existing methods.

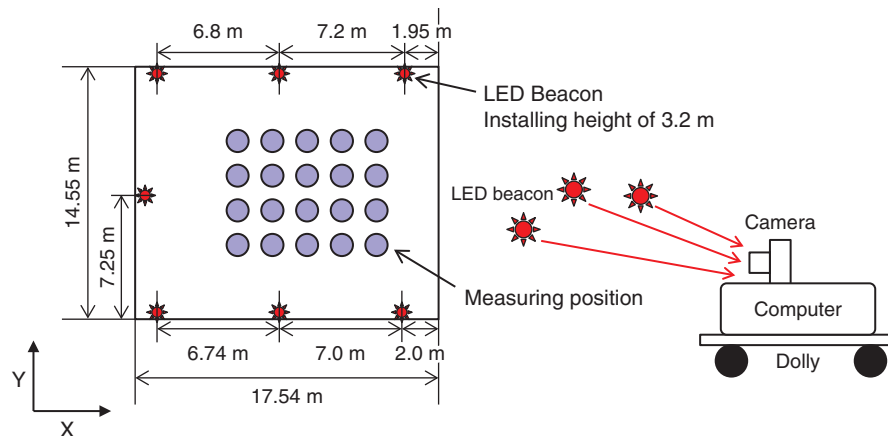


Fig. 12 Position detection

7. Conclusion

In this paper, a communication system receiving an optical signal sent by LED-based light source with a camera, assuming a vehicle-based application, was proposed. By imaging a pulsed light with superimposed digital signals using a lock-in pixel camera, the following performance was achieved:

- (1) Could receive seven optical signals at 640 bps simultaneously
- (2) Could track at a maximum angular velocity of 480 pixels/s
- (3) Could detect distances of 0 to 600 m with an uncertainty of 5.6 m

Additionally, with an assumed application of parking assist, a positioning system using the communication camera was constructed. The following performance was confirmed within a range of approximately 15 m:

- (4) Standard positioning uncertainty of 0.10 m
- (5) Standard angular uncertainty of 0.64°

References

- (1) Heisei 26 nendo Jouhou Tsushin Gijutu no Kenkyu Kaihatsu ni kakaru teianno koubo SIP (Senryakuteki innovation Souzou program) no taishou kadai jidou soukou system, Ministry of Public Management, (2016) (in Japanese)
- (2) 700 MHz tai Koudo Douro Koutsu system Hyoujun Kikaku ARIB STD-T109 1.0 han, Denpa Sangyo kai, (2013) (in Japanese)
- (3) 700 mHz tai wo Mochiita Unten Shien Tsushin system no Jikkenyou guide line, ITS FORUM RC-006 1.0 han, ITS Jouho Tsushin system Suishin Kaigi, (2009) (in Japanese)
- (4) Matsushita, N., Hihara, D., Ushiro, T., Yoshimura, S., Rekimoto, J., Yamamoto, Y.: ID CAM: A Smart Camera for Scene Capturing and ID Recognition, Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality, p. 203-212, (2003)
- (5) Iizuka, N., Kikuchi, M.: Pikapi-camera: Camera wo mochiita kashikou tsuushinn sumaho apuri, Shoumei-gakkai-shi, Vol. 98, No. 10, p. 546-549, (2014) (in Japanese)
- (6) Aoyama, H., Oshima, M.: Line scan sampling for visible light communication: Theory and practice, 2015 IEEE International Conference on Communications, p. 6682-6687, (2015)
- (7) Matsumoto, Y., Nakada, H.: Optical design and receiving characteristics of two-lens visible light communication system, Electronics and Communications in Japan, Vol. 95, No. 10, p. 9-16, (2012)
- (8) Takai, I., Ito, S., Yasutomi, K., Kagawa, K., Andoh, M., Kawahito, S.: LED and CMOS Image Sensor Based Optical Wireless Communication System for Automotive Applications, IEEE Photonics Journal, Vol. 5, Issue 5, (2013)
- (9) Kimachi, A., Ando, S.: Real-Time Phase-Stamp Range Finder Using Correlation Image Sensor, IEEE Sensors Journal, Vol. 9, No. 12, p. 1784-1792, (2009)
- (10) Okuma, T., Kiyokawa, K., Takemura, H., Yokoya, N.: An Augmented Reality System Using a Real-time Vision Based Registration, Proceedings of International Conference on Pattern Recognition, Vol. 2, p. 1226-1229, (1998)
- (11) Aoyama, C., Mineta, K., Okada, Y.: High-resolution range sensor system with stereo camera, JSAE Review, Vol. 21, Issue 3, p. 365-371, (2000)
- (12) Izhal, A. H., Ushinaga, T., Sawada, T., Homma, M., Maeda, Y., Kawahito, S.: A CMOS Time-of-Flight Range Image Sensor with Gates on Field Oxide Structure, IEEE Sensors Journal, Vol. 7, No. 12, p. 141-144, (2005)
- (13) Kawahito, S., Baek, G., Li, Z., Han, S.: CMOS Lock-in Pixel Image Sensors with Lateral Electric Field Control for Time-Resolved Imaging, Proceedings of International Image Sensor Workshop, p. 361-364, (2013)
- (14) Hashida, Y., Ootomo, H., Kuji, Y.: Secondary Surveillance Radar for Air Traffic Control-SSR Mode S, TOSHIBA REVIEW, Vol. 59, No. 2, p. 58-61, (2004)
- (15) Aoyama, C.: Projected Texture Stereo Vision System in Sunlight, Honda R&D Technical Review, Vol. 25, No. 1, p. 132-140
- (16) Aoyama, C.: Camera Calibration by Ray Tracing, Honda R&D Technical Review, Vol. 17, No. 1, p. 138-145
- (17) Nakazawa, Y., Makino, H., Nishimori, K., Wakatsuki, D., Komagata, H.: LED-Tracking and ID-Estimation for Indoor Positioning using Visible Light Communication, 2014 International Conference on Indoor Positioning and Indoor Navigation, (2014)

■ Author ■



Chiaki AOYAMA