

# Rapid and Accurate Time Synchronization using Visible Light for Mobile Sensing

Masanori Sugimoto

*School of Information Science and Technology, Hokkaido University*

Sapporo, Japan, sugi@ist.hokudai.ac.jp

**Abstract**—A novel time-synchronization technique for mobile sensing is proposed. It uses an LED source to transmit a multiplexed linear-chirp optical signal and a rolling-shutter camera to capture a reflected image illuminated by the signal. The proposed technique can estimate the time difference between the start time of the transmitted signal and the shutter release time of the camera both accurately and robustly. Experimental results demonstrated that the technique can achieve time difference estimation errors of  $8.075 \times 10^{-8}$  s at the 90th percentile for a single image.

**Index Terms**—time synchronization, multiplexed chirp signal, rolling-shutter camera

## I. INTRODUCTION

Highly functional and inexpensive image-sensing devices have become available recently, and collecting sensing data from the real world has become much easier, which is accelerating many developments in our data-centered society. For example, the vast and rapid penetration of smartphones with multiple embedded sensors is enabling the seamless data capture of users' daily activity data, which is driving the development of a variety of applications for personal use, such as activity recording and health monitoring. The data collected can also be used for supporting communities through infrastructure applications such as smart cities and smart government [1].

In sensing technologies, spatiotemporal information (i.e., where and when the data is captured) is critical for time-series data analyses that rely on accurate timestamps and must link with contextual and situational information based on accurate localization. For positioning techniques that estimate distance via the propagation time of wireless signals, their performance is affected by the accuracy of time synchronization between transmitters and receivers. The Network Time Protocol (NTP) [2] and the Precision Time Protocol (PTP) [3] which are standard time synchronization techniques for distributed systems on the Internet, and the Flooding Time Synchronization Protocol (FTSP) [4] which is well-known in wireless sensor network communities and its extensions have so far been proposed. These techniques conduct multiple packet exchanges between nodes and may need between several seconds and hours to achieve the required time-synchronization performance. Therefore, they do not always satisfy the demands for real-time systems in dynamic environments.

In this paper, time-synchronization techniques are defined as techniques for estimating time differences between multiple

distributed nodes that each have their own clock. The proposed time-synchronization technique in this paper uses an LED source installed in an indoor environment and a rolling-shutter camera. A modulated optical signal transmitted from the LED illuminates a reflector and the reflected image is photographed by a rolling-shutter camera. The line sensor number having the minimum intensity value is identified from this captured image. By using this number, the time difference between the starting time of the transmitted signal from the LED source and the shutter release timing for this line sensor in the camera can be calculated and used for accurate time synchronization between LED source and camera. A multiplexed linear-chirp signal composed of multiple chirp signals with different sweep frequency bandwidths has been designed using a mathematical model of a rolling-shutter camera working as a frequency filter. As the line sensor number having the minimum intensity converges to the same value when it is detected using chirp signals with different sweep frequency bandwidths, it is possible to achieve accurate and robust time difference estimation.

Experiments to investigate the performance of the proposed technique were conducted to evaluate its performance under a variety of conditions. The experimental parameters varied included the exposure time ratio of the camera, the sweep frequency bandwidth of the multiplexed linear-chirp signal, the distance between the camera and reflector, the type of reflector and the illumination conditions.

The contributions of the paper are summarized as follows.

- A novel time-synchronization technique using an LED source and a rolling-shutter camera is proposed. Multiplexed linear-chirp signals were designed by following the mathematical model of a rolling-shutter camera derived by the author's group [5]. The line sensor number (identified by the minimum intensity of a photographed reflected image) is shown to be represented by a linear relationship involving the exposure time ratio of the camera, the starting time of the transmitted signal, and the shutter release timing of the camera. It is independent of the sweep frequency bandwidths of the constituting chirp signals. By using this linear equation, accurate and robust time difference estimations are available.
- The proposed technique was implemented and evaluated through experiments in real environments. It was found that the technique could achieve time difference estimation errors of  $8.075 \times 10^{-8}$  s at the 90th percentile, using a single image captured from a white paper reflector placed

0.5 m away from the camera with no ambient lighting.

The paper is organized as follows: Section II summarizes previous work related to time-synchronization techniques. Section III describes the details of the proposed technique, including its theoretical aspects and design issues regarding the transmitted signals. Section IV gives the experimental results for the performance evaluation of the proposed technique. Findings indicated through the experimental results are discussed in Section V. Section VI concludes the paper.

## II. RELATED WORK

Many existing time-synchronization techniques and systems use wired or radio-wave communications. These time-synchronization studies are included in this overview to help explain relationships within the proposed technique, although they are implemented in different ways from those using optical signals.

Time-synchronization techniques have been investigated throughout the field of distributed systems. The main problems in realizing time synchronization between multiple nodes are the estimation of the time offset for finding an accurate timestamp, measuring time delays in sending and receiving messages, and detecting clock skew. The NTP [2] and PTP [3] protocols are used to time-synchronize servers and clients on the Internet to achieve submillisecond and submicrosecond accuracy, respectively. However, it takes from tens of seconds to several hours to achieve their target accuracy.

Many methods for WSNs have been proposed. In Reference-Broadcast Synchronization (RBS) [6], neighboring nodes that receive a reference broadcast compare the arrival times according to their local clocks to estimate their time offset and clock skew. One problem with RBS is that the comparison requires many message exchanges ( $(N + 1)/2$  for  $N$  nodes). In delay-measurement time synchronization [7], a one-time broadcast from a master node allows slave nodes to time synchronize by estimating the time delay at each slave node. In the timing-sync protocol for sensor networks [8], a root node works as a master clock node and extends a hierarchical structure by exchanging timestamps between a parent and its child nodes to identify the clock drift and propagation delay between them. FTSP [4] sends periodical flooding messages from a master node. A slave node conducts regression calculations by employing multiple pairs of global timestamps from the master node and its own local timestamps. Glossy [9] implements a time-synchronization technique similar to that of FTSP. To achieve successful message decoding without capture effects, it sets periodical time slots for message flooding that are exclusively separated from other application tasks executed on each node. In FTSP, clock skews between master and local nodes increase exponentially with the number of nodes. PulseSync [10] alleviates this problem and has been demonstrated as reducing the clock skews to the order of the square root of the number of nodes. The time-of-flight aware time-synchronization protocol [11] achieves submicrosecond accuracy over 22 hops by conducting propagation delay compensation. CESP [12] was proposed as a method for estimating

time offset and drift between master and slave nodes, by receiving packets from reference nodes to reduce communication overhead and energy consumption. Shi et al. [13] extended FTSP from one-way to two-way message exchange with maximum-likelihood estimation to generate time skew/offset compensation. A problem with average-consensus-based time-synchronization algorithms [14] is longer convergence time required to achieve a common value among nodes, caused by the many iterative message exchanges between them. MACTS [15] generates virtual connections between multihop nodes, aiming to reduce this convergence time.

Zhong et al. [16] proposed On-Demand Time Synchronization (ODS), which can adjust clock calibration intervals for desired levels of accuracy. After the demand for a particular offset estimation error is set, ODS adjusts its parameters to satisfy the demand with a given confidence probability. Virtual high-resolution time [17] used high-speed and low-speed clocks to achieve high-resolution time synchronization with low power consumption by coordinating the use of the two clocks. Time delays, including timestamps when sending and receiving messages, are modeled as exponential random variables. To achieve time synchronization in a computationally efficient manner, a low-complexity maximum-likelihood estimator has been proposed [18]. A general model for estimating clock skew and time offsets using the Kalman Filter method has also been developed [19].

Radio-wave-based techniques that run on WSNs [4], [9]–[11] can achieve submicrosecond synchronization but have to be implemented using real-time operating systems, which are yet to be adopted for current smartphones and tablet PCs. This limitation makes it much more difficult to obtain accurate and precise timestamps using commercially off-the-shelf (COTS) mobile devices than it is using WSNs. When a target application implements indoor positioning in mobile settings, it is desirable to complete the time synchronization as rapidly as possible. Radio-wave-based techniques conduct repeated message exchanges between nodes to estimate their timestamps accurately, but such frequent message exchanges may induce traffic congestion or require long synchronization times (several seconds to tens of minutes).

In an alternative approach, an ultrasound time-synchronization technique has been proposed [20]. By calculating the location of the smartphone using time-difference-of-arrival trilateration, the time offset for sending a signal from the speakers time-synchronized with a master clock can be estimated. If a microphone array using multiple smartphones is used to estimate the direction of arrival of sounds, all microphones in the array must be time-synchronized. Dia [21] adopted two-level synchronization, whereby synchronization between one master smartphone and the other smartphones, together with synchronization between the CPU clock and the audio I/O clock in each smartphone, was conducted using a least-squares regression method. BeepBeep [22] conducted acoustic ranging for a pair of smartphones by estimating the signal propagation time between the smartphones. Studies such as this do not clearly

state how long the synchronization takes, but completing it rapidly would appear to be difficult. One example [21] conducts regression calculations using multiple timestamp data obtained through message exchanges and others [20], [22] employ specific time intervals between the signal emissions from each smartphone to avoid reverberations, thereby making rapid synchronization difficult.

Regarding time-synchronization techniques using optical signals, FLIGHT [23] was implemented using a fluorescent lamp and the 10-kHz-sampling built-in light sensor on a WSN sensor node called TelosB. The stable nature of the alternating-current frequency enabled time synchronization within tens of microseconds. Alternatively [24] uses a single LED that can be used for transmitting and receiving optical signals. When the LED is switched to receiver mode, it can be used to measure the differences in the charged voltages between two time slots to achieve time synchronization.

Some techniques using a camera have been proposed for illumination-based time synchronization. PSync [25] used an LED illumination source as a transmitter of De Bruijn sequences and a photo detector as a receiver for rapid and energy-efficient time synchronization. Akiyama et al. [26] and Sugimoto et al. [5] proposed techniques using a camera to capture modulated optical signals emitted from LED sources for acoustic localization. To achieve rapid execution of an algorithm based on a phase-locked loop, use of a high-speed (1,000 fps) camera has been proposed [27]. SocialSync [28] can time synchronize multiple smartphones for 3D construction and scene-depth estimation. This technique can estimate the delay between a camera I/O event and a timestamp given by the application software to receive an image frame with millisecond accuracy.

Our proposed technique can be implemented using COTS mobile devices without any additional hardware, and its rapid time synchronization can be achieved with submicrosecond accuracy, which is not possible with existing camera-based systems or techniques.

### III. PROPOSED TECHNIQUE

This section describes issues regarding the proposed technique.

#### A. Mathematical Model of a Rolling-shutter Camera

Aiming for low energy consumption, smartphone built-in cameras implement a rolling-shutter technique, whereby the incoming optical 2D image is treated as an array of scan lines, with the overall image being captured as a sequence of line-by-line scans from each line sensor. A mathematical model of the rolling-shutter camera is briefly described in this section (see [29] for details).

#### B. Design of the Transmit Signal

Let  $s(t)$  denote a periodical optical signal emitted from an LED source and let  $f_C$ ,  $T(=1/f_C)$ ,  $\eta$ , and  $L$  denote the frame rate, frame period, exposure time ratio, and total number of line sensors for a camera, respectively. Here,  $f_C$  is regarded

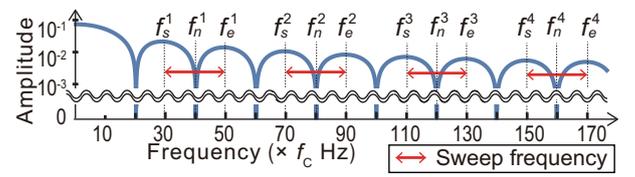


Fig. 1. A rolling-shutter camera as a frequency filter ( $\eta=0.05$ ).

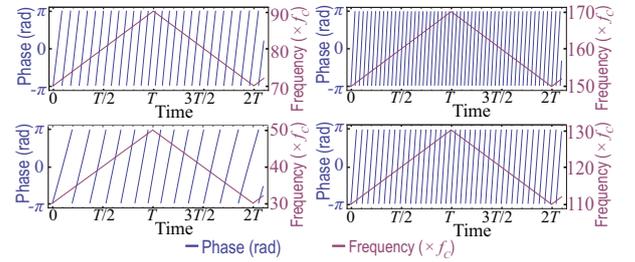


Fig. 2. Frequencies and phases of a multiplexed linear-chirp signal ( $\eta=0.05$ ,  $N_f=4$ ).

as the fundamental frequency of  $s(t)$  when the readout time of the line sensors is not considered and  $\delta T$  ( $0.0 \leq \delta < 1.0$ ) is the time difference between the starting time of  $s(t)$  and the shutter release time of the first line sensor. The intensity value  $r(\tau_l)$  at the  $l$ -th line sensor is given by Equation (1).

$$r(\tau_l) = \frac{1}{T} \int_0^{\eta T} s(t + \delta T + \frac{l}{L}T) dt \quad (1)$$

Note that  $\tau_l = \delta T + lT/L$  represents the time difference between the starting time of  $s(t)$  and shutter release time of the  $l$ -th line sensor. Let  $S_k$  and  $R_k(l)$  ( $k = 0, \pm 1, \pm 2, \dots$ ) denote the Fourier transforms of  $s(t)$  and  $r(\tau_l)$ , respectively. The value for  $R_k(l)/S_k$ , their spectrum ratio, is then given by Equation (2) and is illustrated in Figure 1.

$$\frac{R_k(l)}{S_k} = \eta e^{jk\eta\pi(1+2l/L)} \text{sinc}(k\eta\pi) \quad (2)$$

Here,  $k$  means the frequency order of the camera frame rate  $f_C$  and  $\text{sinc}(x)$  is defined as  $\text{sinc}(x) = \sin(x)/x$ .

1) *Multiplexed Linear-chirp Signal and its Characteristics*: Figure 1 shows that there are multiple frequencies  $f_n^i$  ( $i = 1, 2, \dots$ ) for which  $R_k(l)/S_k = 0$  (null) holds. From Equation (2),  $k\eta = m_i$  ( $m_i$ : integer other than zero) is obtained because  $\text{sinc}(k\eta\pi) = 0$  holds. A multiplexed linear-chirp signal represented by Equation (3) and transmitted from an LED source is composed of a chirp signal whose sweep frequency band is  $[f_s^i - f_e^i]$ , including  $f_n^i = m_i/\eta f_C$ , with  $f_s^i$  and  $f_e^i$  being the nearest spectrum peaks to  $f_n^i$  (Figure 1).

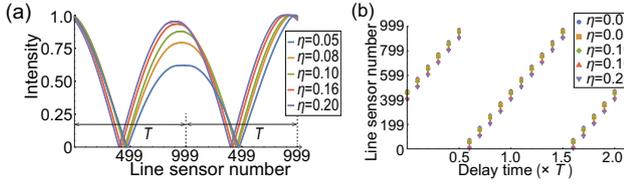


Fig. 3. (a) Intensity values for the range of line sensors (normalized by the peak intensity value at each  $\eta$  setting,  $L=1,000$ ,  $\delta T=0.0$ ) and (b) Relation between  $\delta T$  and  $l_{min}$  ( $L=1,000$ ).

$s(t) =$

$$s(t) = \begin{cases} \sum_{i=1}^{N_f} \exp \left[ j2\pi \left( \frac{f_e^i + f_s^i}{2T} \text{Quotient}[x, y] + f_s^i \text{Mod}[t, 2T] + \frac{f_e^i - f_s^i}{2T} \text{Mod}[t, 2T]^2 \right) \right], & (0 \leq \text{Mod}[t, 2T] < T) \\ \sum_{i=1}^{N_f} \exp \left[ j2\pi \left( \frac{f_e^i + f_s^i}{2T} \text{Quotient}[x, y] + f_s^i (\text{Mod}[t, 2T] - T) + \frac{f_e^i - f_s^i}{2T} (\text{Mod}[t, 2T] - T)^2 \right) \right], & (T \leq \text{Mod}[t, 2T] < 2T) \end{cases} \quad (3)$$

Here,  $N_f$  ( $i = 1, 2, \dots, N_f$ ),  $\text{Quotient}[x, y]$  and  $\text{Mod}[x, y]$  are the number of chirp signals composing  $s(t)$ , the integer part of the quotient, and the remainder after dividing dividend  $x$  by divisor  $y$ , respectively.

The signal  $s(t)$  is continuous, with a period of  $2T$ , and is composed of a pair of up-chirp and down-chirp signals, each lasting  $T$ . Because their frequencies and phases change continuously, as shown in Figure 2, the intensity values obtained by the camera (given by Equation (1)) will also change continuously. The line sensor number having the minimum intensity value when receiving either up-chirp or down-chirp signals is denoted  $l_{min}$  ( $0 \leq l_{min} \leq L - 1$ ) and is the same value in all sweep frequency bands  $[f_s^i - f_e^i]$ . Figure 3 (a) shows the intensity value for the range of line sensor numbers when  $\delta T$  is set to 0.  $l_{min} = \arg \min_l r(\tau_l)$  is obtained by Equation (4).

This was confirmed numerically through computer simulations and preliminary experiments in real environments, because  $l_{min}$  is found as a solution to  $\partial r(\tau_l) / \partial l = 0$ , which includes Fresnel integrals and cannot be solved analytically. The time difference  $\delta T$  is calculated by Equation (5).

$$l_{min} = \text{Mod}(0.5(1 - \eta) + \delta, 1.0)L \quad (4)$$

$$\delta T = \text{Mod} \left( \frac{\arg \min_l r(\tau_l)}{L} - 0.5(1 - \eta), 1.0 \right) T \quad (5)$$

Figure 3 (b) shows the relation between  $\delta T$  and  $l_{min}$ , confirming that  $\delta T$  can be estimated by using Equation (5) and  $l_{min}$ . The period of  $s(t)$  is  $2T$  and the values of  $\delta$  obtained by the proposed technique are found in the range  $0.0 \leq \delta < 1.0$ . Therefore,  $\delta T$  can be determined as the time difference from the starting time of the up-chirp or down-chirp signal.

The value for  $l_{min}$  obtained by the proposed technique does not depend on the sweep frequency bandwidths of the

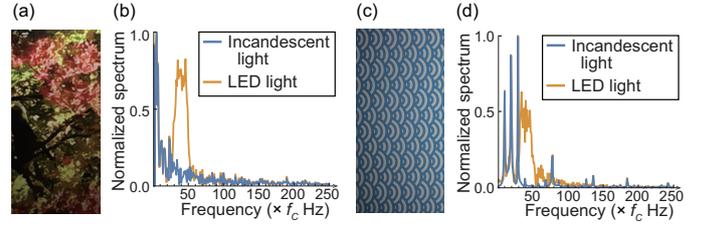


Fig. 4. Reflected images and their spatial frequency spectra: (a) nature image, (b) frequency spectra for image(a), (c) pattern image, and (d) those for image(c).

chirp signals. It is therefore possible to estimate the time difference accurately and robustly from a range of  $l_{min}$  values for multiple chirp signals with different sweep frequency bandwidths. Accurate estimation is further enhanced by using quadratic function interpolation with the intensity values of  $l_{min}$  and those of the neighboring sensors to estimate  $l_{min}$  for intermediate points.

2) *Sweep Frequency*: To take full advantage of the features of the rolling-shutter camera, the transmitted optical signal should be received by using all its line sensors. This implies that the transmitted signal should not be captured as direct light from an LED but as an indirect image via a reflector. To determine the sweep frequency bandwidths of transmitted chirp signals, two issues should be considered. First, the frequencies of the signals must exceed the critical fusion frequency to avoid users' perception of flickering [30], and be  $L/2 f_C$  or below because of the sampling theorem. Second, any interference with the spatial frequency spectra of the reflector by the chosen frequency bandwidths of the chirp signals should be attenuated as much as possible.

Figures 4 (a) and 4 (c) show the captured images (exposure time ratio  $\eta = 0.10$ ) of a photograph of red leaves (the "nature image" in (a)) and a traditional Japanese blue-ocean-wave-pattern (the "pattern image" in (c)) illuminated by a chirp signal (sweep frequency in the range  $[30.0f_C - 50.0f_C]$ ), which were used as the reflected images. Figures 4 (b) and 4 (d) show the frequency spectra of the images shown in Figures 4 (a) and 4 (c). The distributions of the frequency spectra when illuminated by an incandescent light (the blue line in Figures 4 (b) and 4 (d)) can be approximated to those of the original image because the distribution of the frequency spectra for this light is regarded as uniform. From the figure, note that that the nature image and pattern image both have strong spectra in the low-frequency bandwidths (lower than  $30f_C$  Hz) and the pattern image has strong peaks in the sweep frequency bandwidth  $[30.0f_C - 50.0f_C]$  and above. Based on this observation, the lowest frequency for the chirp signal was set as  $30.0f_C$  Hz. The influence of the frequency spectra in the high-frequency bandwidths shown for the pattern image is discussed in Sections IV-B3.

### C. Time Difference Estimation Algorithm

The time difference  $\delta T$  is calculated via the following procedure.

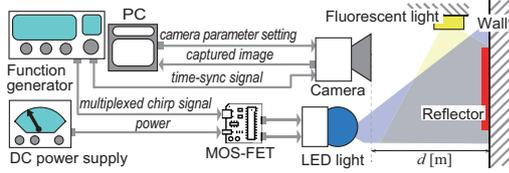


Fig. 5. Experimental environment.

- 1) The sweep frequency bandwidths for the multiplexed linear-chirp signal are determined from the value of the exposure time ratio  $\eta$  (see Table II for details). That is,  $[f_s^i, f_e^i]$  is the narrowest frequency bandwidth including  $f_n^i \in [f_s^i, f_e^i]$ , where  $f_s^i$  and  $f_e^i$  denote the peak frequency spectra and  $f_n^i$  denotes the null spectrum (see Equation (2) and Figure 1).
- 2) The multiplexed linear-chirp signal represented by Equation (3) is then transmitted from the LED source.
- 3) The RGB image  $I_{RGB}(i, j)$  ( $0 \leq i < M, 0 \leq j < L$ ) from the reflector illuminated by the LED source is captured by the rolling-shutter camera ( $L \times M$  pixels) and cropped to decrease its size to  $L \times M_t$  pixels by selecting a rectangular area determined by  $((M - M_t)/2 + 1, 0)$ ,  $((M - M_t)/2 + 1, L)$ ,  $((M + M_t)/2, L)$ , and  $((M + M_t)/2, 0)$ . The RGB image is converted to an intensity image  $I_Y(u, v)$  by using the equation  $y = 0.299r + 0.587g + 0.114b$ , where  $r$ ,  $g$ , and  $b$  are the R, G, and B values for the RGB image and  $y$  is the intensity value at pixel  $(u, v)$  ( $0 \leq u < M_t, 0 \leq v < L$ ).
- 4) The mean intensity value for the  $l$ -th line sensor  $i_l$  is calculated using  $i_l = \sum_{u=0}^{M_t-1} I_Y(u, l)$  to obtain the one-dimensional (1D) intensity vector  $\mathbf{I} = (i_0, i_1, \dots, i_{L-1})$ . A fast Fourier transform is applied to  $\mathbf{I}$  and the spectra of  $\mathbf{I}$  are separated into the frequency bandwidth for each chirp signal contributing to the multiplexed linear-chirp signal.
- 5) The spectra in each frequency bandwidth are converted to their analytic signals. The line sensor number  $l_{min}$ , which has the minimum magnitude of the signal, and the magnitudes of its neighboring line sensors are obtained. By applying quadratic function interpolation, intermediate line sensor numbers are calculated. From Equation (5), the time difference  $\delta T$  is found. If the frequency distribution for the reflector is available beforehand,  $l_{min}$  can be calculated using only the chirp signals that do not have strong spatial frequency spectra from the reflector in their sweep frequency bandwidths.

#### IV. EXPERIMENTS

This section shows experimental results of the proposed system.

##### A. Overview

The experimental environment used to evaluate the proposed technique is shown in Figure 5. The multiplexed linear-chirp signal was generated by a function generator (NF Corporation WF1948), which was connected to an LED source (W-LITE

DLFL-001) via a switching power MOS-FET connected to a DC power supply. The purpose of the experiment was to investigate the performance of the proposed method for time difference estimation. The estimation involves the start time of the chirp signal and the shutter release timing of the camera. Being able to release the shutter at an accurate time difference  $\delta_{true}T$  and set camera parameters flexibly is critical to the evaluation of the proposed technique, and a high-functional rolling-shutter camera (Point Grey/Teledyne Flir FL3-U3-13S2-C) was therefore used instead of a smartphone's built-in camera. The camera received a trigger signal for shutter release directly from the function generator via its general-purpose I/O interface. The image from the wall-mounted reflector was captured by the camera and transferred via its USB 3.0 interface to a PC, which conducted the calculations described in Section III-C to estimate the time difference  $\delta T$ . The proposed technique was evaluated by comparing  $\delta T$  with  $\delta_{true}T$ . From inspection of the camera, the maximum exposure time of the line sensors, excluding their readout time, was found to be 8.028 ms for a frame rate of 50 Hz. Therefore,  $f_C$ , the fundamental frequency of the transmitted signal in this experiment, was set to  $1/8.028 = 124.56$  Hz. The resolution of the camera was  $1024 \times 1328$  pixels and it exposes two neighboring line sensors simultaneously. The total number of line sensors ( $L$ ) was therefore  $1024/2 = 512$ .

As described in Section III-C, the captured image was converted to a 1D vector by averaging the intensity values for the same line sensor. (When many values can be averaged, the noise and the influence of the reflector can be suppressed.) After considering computational complexity and the effects of lens vignetting, a  $1024 \times 500$ -pixel image whose center coincided with the camera optical axis was cropped from the captured image. The sweep frequency bandwidths of the transmitted signals, composed of four pairs of up-chirp and down-chirp for each exposure time ratio  $\eta$ , are given in Table II.

The experimental parameters  $\eta$ ,  $ilc$  (illumination condition),  $d$  (distance between the camera and reflector) and  $sbj$  (reflector) were set to a variety of values in five experiments, as shown in Table I. The correct values for the time difference  $\delta_{true}T$  (measured using the trigger signal from the function generator) were set as  $\delta_{true}T = 0.1 \times i T$  ( $i = 0, 1, \dots, 19$ ). The time difference estimation was repeated 200 times for each experimental parameter setting, giving  $5 \times 3 \times 3 \times 2 \times 20 \times 200 = 360,000$  image captures for the experiments.

##### B. Experimental Results

1) *Experiment 1: Exposure time ratio:* Figure 6 shows the cumulative distribution function (CDF) values, for each exposure time ratio, of the time difference estimation errors produced using our proposed technique and Figure 7 shows their mean and standard deviation values. Table III shows the 90th percentile errors obtained using single chirp signals for different sweep frequency bandwidths. The best estimation performance results for each exposure time are shown in bold. This experiment demonstrates that the proposed technique can

TABLE I  
 PARAMETER SETTINGS FOR EACH EXPERIMENT.

| Experiment   | Exposure time ratio ( $\eta$ ) | Illumination condition ( $ilc$ )              | Distance ( $d$ [m]) | Reflector ( $sbj$ )                      |
|--------------|--------------------------------|---|---------------------|--|
| Experiment 1 | 0.01, 0.05, 0.08, 0.16, 0.2    | No ambient light                              | 0.5                 | White paper                              |
| Experiment 2 | 0.01, 0.05, 0.08, 0.16, 0.2    | No ambient light                              | 0.5, 1.0            | White paper                              |
| Experiment 3 | 0.01, 0.05, 0.08, 0.16, 0.2    | No ambient light                              | 0.5                 | White paper, Nature image, Pattern image |
| Experiment 4 | 0.01, 0.05, 0.08, 0.16, 0.2    | No ambient light, Fluorescent light, Sunlight | 0.5                 | White paper                              |

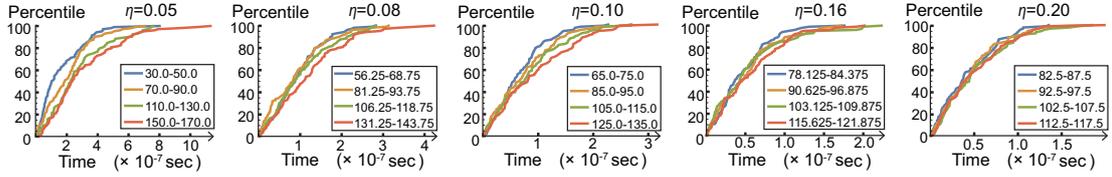


Fig. 6. CDFs of time difference estimation errors for different exposure time ratios.

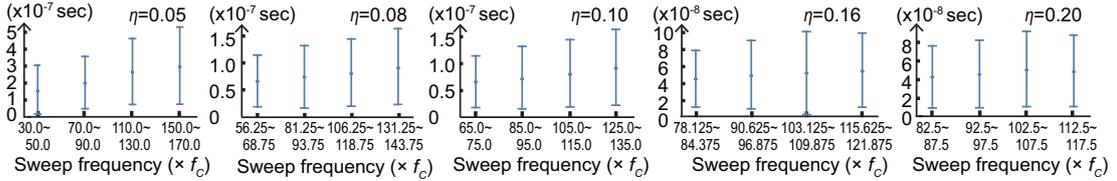


Fig. 7. Means and standard deviations of time difference estimation errors for different exposure time ratios.

 TABLE II  
 SWEEP FREQUENCY BANDWIDTHS [ $f_s^i - f_e^i$ ] ( $\times f_C$ ) OF CHIRP SIGNALS FOR DIFFERENT EXPOSURE TIME RATIOS  $\eta$ .

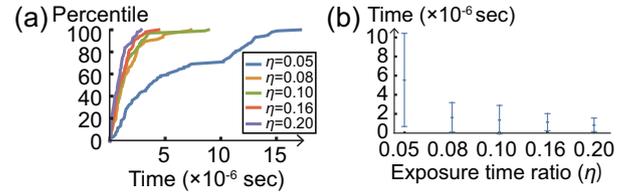
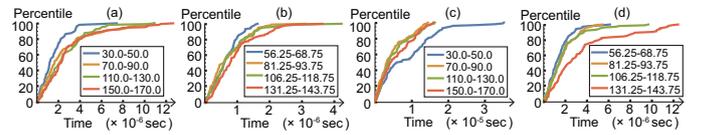
| $\eta$ | $f_s^1 - f_e^1$ | $f_s^2 - f_e^2$ | $f_s^3 - f_e^3$ | $f_s^4 - f_e^4$ |
|--------|-----------------|-----------------|-----------------|-----------------|
| 0.05   | 30.0-50.0       | 70.0-90.0       | 110.0-130.0     | 150.0-170.0     |
| 0.08   | 56.25-68.75     | 81.25-93.75     | 106.25-118.75   | 131.25-143.75   |
| 0.10   | 65.0-75.0       | 85.0-95.0       | 105.0-115.0     | 125.0-135.0     |
| 0.16   | 78.125-84.375   | 90.625-96.875   | 103.125-109.375 | 115.625-121.875 |
| 0.20   | 82.5-87.5       | 92.5-97.5       | 102.5-107.5     | 112.5-117.5     |

 TABLE III  
 90TH PERCENTILE ERRORS FOR DIFFERENT EXPOSURE TIME RATIOS.

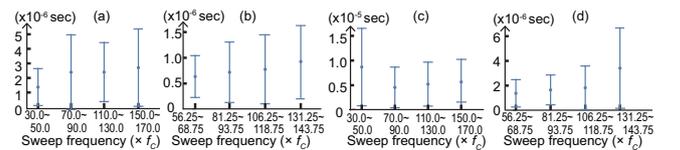
| Frequency       | $\eta = 0.05$          | $\eta = 0.08$          | $\eta = 0.10$          | $\eta = 0.16$          | $\eta = 0.20$          |
|-----------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| $f_s^1 - f_e^1$ | $3.336 \times 10^{-7}$ | $1.607 \times 10^{-7}$ | $1.128 \times 10^{-7}$ | $8.075 \times 10^{-8}$ | $1.012 \times 10^{-7}$ |
| $f_s^2 - f_e^2$ | $3.418 \times 10^{-7}$ | $1.593 \times 10^{-7}$ | $1.322 \times 10^{-7}$ | $9.285 \times 10^{-8}$ | $1.112 \times 10^{-7}$ |
| $f_s^3 - f_e^3$ | $4.791 \times 10^{-7}$ | $1.898 \times 10^{-7}$ | $1.544 \times 10^{-7}$ | $9.650 \times 10^{-8}$ | $1.273 \times 10^{-7}$ |
| $f_s^4 - f_e^4$ | $5.623 \times 10^{-7}$ | $2.157 \times 10^{-7}$ | $1.753 \times 10^{-7}$ | $1.035 \times 10^{-7}$ | $1.164 \times 10^{-7}$ |

achieve an  $8.075 \times 10^{-8}$  error at the 90th percentile. Significant performance differences for the various sweep frequency bandwidths were not observed, but the best performance was achieved by using the sweep frequency bandwidth [ $f_s^1 - f_e^1$ ].

2) *Experiment 2: Distance*: Images from the reflector were captured by the camera for distances of  $d=0.5$  m and  $d=1.0$  m between reflector and camera. The time difference estimation was conducted using the multiplexed linear-chirp signal. The results obtained using the chirp signal for [ $f_s^1 - f_e^1$ ] are shown in Figures 8 (a) and (b). (Because of page limitations and similarities to the results of Experiment 1, the results for other chirp signal frequencies are not shown.) These two figures show that the 90th percentile errors at the  $d=1.0$  m setting are  $1.219 \times 10^{-5}$  s ( $\eta=0.05$ ),  $2.794 \times 10^{-6}$  s ( $\eta=0.08$ ),  $2.640 \times 10^{-6}$  s ( $\eta=0.10$ ),  $2.108 \times 10^{-6}$  s ( $\eta=0.16$ ) and  $2.383 \times 10^{-6}$  s ( $\eta=0.20$ ). Note that the performance deterioration is most significant for the  $\eta=0.05$  setting than for the other settings.


 Fig. 8. Time difference estimation errors at the  $d=1.0$  m setting: (a) CDFs and (b) Means and standard deviations

 Fig. 9. CDFs using different reflectors: (a) nature image ( $\eta=0.05$ ), (b) nature image ( $\eta=0.08$ ), (c) pattern image ( $\eta=0.05$ ), and (d) pattern image ( $\eta=0.08$ ).

3) *Experiment 3: Reflector*: In Figure 4 (d), note that the pattern image has strong spectra around the 78th and 136th frequency orders of  $f_C$ . An experiment was therefore conducted to investigate the influence of the spatial frequencies associated with the reflector. Figures IV-B2 and 10 show the CDFs, means, and standard deviations of the time dif-


 Fig. 10. Means and standard deviations using different reflectors: (a) nature image ( $\eta = 0.05$ ), (b) nature image ( $\eta=0.08$ ), (c) pattern image ( $\eta=0.05$ ), and (d) pattern image ( $\eta=0.08$ ).

ference estimation errors by using multiplexed linear-chirp signals ( $\eta = 0.05$  and  $0.08$ ) for the  $sbj$ ="nature image" and  $sbj$ ="pattern image" settings. The 90th percentile error, mean, and standard deviation obtained by the chirp signal with a sweep frequency  $[70.0f_C - 90.0f_C]$  ( $\eta=0.05$ ) (including the 78th frequency order) for the  $sbj$ ="pattern image" settings were  $1.808 \times 10^{-5}$  s,  $8.681 \times 10^{-6}$  s, and  $7.836 \times 10^{-6}$  s, respectively. The results for the chirp signal with a sweep frequency  $[131.25f_C - 142.75f_C]$  ( $\eta=0.08$ ) (including the 136th frequency order) for the  $sbj$ ="pattern image" settings were  $7.753 \times 10^{-6}$  s,  $3.411 \times 10^{-6}$  s, and  $3.252 \times 10^{-6}$  s, respectively. From Figures IV-B2 and 10, note the deterioration in time difference estimation performance using a chirp signal with strong spatial frequency spectra from the "pattern image" in its sweep frequency bandwidth. This deterioration is not apparent using a chirp signal with minimal spatial frequency spectra from the "nature image" in its sweep frequency bandwidth.

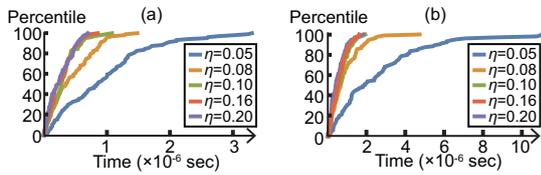


Fig. 11. CDFs of time difference estimation errors for different illumination conditions: (a) fluorescent light and (b) sunlight.

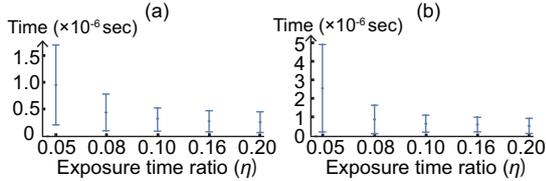


Fig. 12. Means and standard deviations of time difference estimation errors for different illumination conditions: (a) fluorescent light and (b) sunlight.

4) *Experiment 4: Illumination*: Experimentation with different illumination settings were conducted and the results obtained by using a chirp signal having the sweep frequency bandwidth  $[f_s^1 - f_e^1]$  are shown in the same way as for Experiment 2. Figure 11 demonstrates that the 90th percentile errors for the  $ilc$ ="Fluorescent light" setting were  $1.668 \times 10^{-6}$  s ( $\eta=0.05$ ),  $8.525 \times 10^{-7}$  s ( $\eta=0.08$ ),  $5.429 \times 10^{-7}$  s ( $\eta=0.10$ ),  $4.890 \times 10^{-7}$  s ( $\eta=0.16$ ), and  $5.905 \times 10^{-7}$  s ( $\eta=0.20$ ), respectively. The figure also shows that the results for the  $ilc$ ="Sunlight" setting are  $4.676 \times 10^{-6}$  s ( $\eta=0.05$ ),  $1.530 \times 10^{-6}$  s ( $\eta=0.08$ ),  $1.122 \times 10^{-6}$  s ( $\eta=0.10$ ),  $1.044 \times 10^{-6}$  s ( $\eta=0.16$ ), and  $1.210 \times 10^{-6}$  s ( $\eta=0.20$ ), respectively. Note that the most significant deterioration is observed for the  $\eta=0.05$  setting than for the other settings.

## V. DISCUSSION

This section discusses findings and limitations regarding the proposed technique.

### A. Findings Derived from the Experimental Results

Findings from the experimental results are discussed as follows.

- The results of Experiment 1 imply that the larger the exposure time ratio  $\eta$ , the better the estimation performance of time differences, which is related to the improvement in signal-to-noise (S/N) ratios. Performance differences between the various chirp signals are not significant for the  $sub$ ="white paper" setting, which does not affect the spectrum distributions in their frequency bandwidths.
- Experiment 2 demonstrates that the 90th percentile errors using the chirp signal having the sweep frequency bandwidth  $f_s^1 - f_e^1$  for the  $d = 1.0$  setting are worse than those for the  $d = 0.5$  m setting, as shown in Table III and are also related to S/N ratios. Therefore, estimation performance requirements can be optimized by adjusting the power of transmitted signals and/or the distance to the reflector.
- From the results of Experiment 3, performance deterioration is observed when strong spatial frequency spectra associated with the reflector exist in the sweep frequency bandwidths of the chirp signals. If the spectrum distribution of the reflector is known beforehand, it should be possible to retain the estimation performance by selecting chirp signals having appropriate sweep frequency bandwidths and/or by changing exposure time ratios to avoid the spectrum interference. For example, if the reflector shown in Figure 4(b) is used, the estimation performance would best be retained by using chirp signals having sweep frequency bandwidths other than  $[f_s^1 - f_e^1]$  ( $\eta = 0.05$ ) or  $[f_s^4 - f_e^4]$  ( $\eta = 0.08$ ) and/or by adjusting the exposure time ratio. Figures IV-B2 (c) and IV-B2(d) show that the CDFs using a chirp signal strongly affected by the reflector spectra are clearly different from those less affected. Therefore, if the spectrum distribution of the reflector is unknown, it would be worth investigating the effectiveness of a method that compared the estimation results obtained by individual chirp signals and then eliminating those results that are significantly different from the other results as "outliers" caused by reflectors.
- Experiment 4 demonstrated that the time difference estimation performance deteriorates under the  $ilc$ ="Fluorescent light" and "Sunlight" settings, in comparison with that for the  $ilc$ ="No ambient light" setting, as shown in Table III. This is caused by their poorer S/N ratios. The estimation results under the  $ilc$ ="Sunlight" setting were worse than for  $ilc$ ="Fluorescent light" setting, which might again relate to spectral interference with the chirp signals in their various sweep frequency bandwidths. To retain the performance of the proposed technique, therefore, spectrum interference between the transmitted signal and the ambient lighting should be avoided where possible.

### B. Limitations

The limitations of this study from a practical point of view can be summarized as follows.

- With many commercially available off-the-shelf cameras, it is not possible to set their camera parameters to arbitrary values.

trary values, but some setting to predetermined discrete values is allowed. For example, the camera used in this study does not allow setting its exposure time ratio to an arbitrary value. Therefore, it may not be possible to set the camera parameters to the theoretically derived optimum values suggested by Equation (2), which would cause the performance deterioration.

- In the author's previous study [5], it was noted that the camera frame rate may not be stable, with clock drift being observed. Therefore, to retain the performance of the time difference estimation, the frame rate of the camera should be estimated and updated periodically.
- The experimental evaluations were conducted in completely static and stable settings. However, in more realistic situations, the camera and reflector might well be mobile. In future work, the performance of the proposed technique in mobile or unstable environments will be investigated.

## VI. CONCLUSIONS

A novel time-synchronization technique is proposed in this paper. The proposed technique uses an LED source transmitting multiplexed linear-chirp signals to a reflector and a rolling-shutter camera that photographs the reflected image. It can estimate the time difference between the start time of the transmitted signal and the shutter release time of the camera both accurately and robustly. Our experiments demonstrated that the proposed technique can achieve  $8.075 \times 10^{-8}$ -s time difference estimation errors at the 90th percentile by using a single image captured 0.5 m away from the reflector. Implementing and evaluating the proposed technique using a smartphone's built-in camera are the next steps in this study.

## REFERENCES

- [1] R. Khatoun and S. Zeadally, "Smart Cities: Concepts, Architectures, Research Opportunities," *Commun. ACM*, vol. 59, no. 8, pp. 46–57, 2016.
- [2] D. L. Mills, "Internet Time Synchronization: The Network Time Protocol," *IEEE Trans. Commun.*, vol. 39, no. 10, pp. 1482–1493, 1991.
- [3] IEEE 1588-2019, "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems," <https://standards.ieee.org/ieee/1588/6825/> (retrieved: Oct 6th, 2022), 2020.
- [4] M. Maróti, B. Kusy, G. Simon, and A. Lédeczi, "The Flooding Time Synchronization Protocol," in *Proc. of SenSys'04*, Baltimore, MD, 2004, pp. 39–49.
- [5] M. Sugimoto, H. Kumaki, T. Akiyama, and H. Hashizume, "Optimally Modulated Illumination for Rapid and Accurate Time Synchronization," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 505–516, 2017.
- [6] J. Elson, L. Girod, and D. Estrin, "Fine-grained Network Time Synchronization using Reference Broadcasts," in *Proc. of OSDI'02*, Boston, MA, 2002, pp. 147–156.
- [7] S. Ping, "Delay Measurement Time Synchronization for Wireless Sensor Networks," Intel Research Berkeley Lab., Tech. Rep. IRB-TR-03-013, 2003.
- [8] S. Ganeriwala, R. Kumar, and M. B. Srivastava, "Timing-sync Protocol for Sensor Networks," in *Proc. of SenSys'03*, Los Angeles, CA, 2003, pp. 138–149.
- [9] F. Ferrari, M. Zimmerling, L. Thiele, and O. Saukh, "Efficient Network Flooding and Time Synchronization with Glossy," in *Proc. of IPSN 2011*, Chicago, IL, 2011, pp. 73–84.
- [10] C. Lenzen, P. Sommer, and R. P. Wattenhofer, "Optimal Clock Synchronization in Networks," in *Proc. of SenSys'09*, Berkeley, CA, 2009, pp. 225–238.
- [11] R. Lim, B. Maag, and L. Thiele, "Time-of-Flight Aware Time Synchronization for Wireless Embedded Systems," in *Proc. of EWSN'16*, Graz, Austria, 2016, pp. 149–158.
- [12] F. Gong and M. L. Sichitiu, "CESP: A Low-Power High-Accuracy Time Synchronization Protocol," *IEEE Trans. Vehicular Tech.*, vol. 65, no. 4, pp. 2387–2396, 2016.
- [13] F. Shi, S. X. Yang, X. Tuo, L. Ran, and Y. Huang, "A Novel Rapid-Flooding Approach With Real-Time Delay Compensation for Wireless-Sensor Network Time Synchronization," *IEEE Trans. Cybern.*, vol. 52, no. 3, pp. 1415–1428, 2022.
- [14] Q. Li and D. Rus, "Global Clock Synchronization in Sensor Networks," *IEEE Transactions on Computers*, vol. 55, no. 2, pp. 214–226, 2006.
- [15] F. Shi, X. Tuo, L. Ran, Z. Ren, and S. Yang, "Fast Convergence Time Synchronization in Wireless Sensor Networks Based on Average Consensus," *IEEE Trans. Industr. Inform.*, vol. 16, no. 2, pp. 1120–1129, 2020.
- [16] Z. Zhong, P. Chen, and T. He, "On-demand Time Synchronization with Predictable Accuracy," in *Proc. of INFOCOM 2011*, Shanghai, China, 2011, pp. 2480–2488.
- [17] T. Schmid, P. Dutta, and M. B. Srivastava, "High-Resolution, Low-Power Time Synchronization an Oxymoron No More," in *Proc. of IPSN 2010*, Stockholm, Sweden, 2010, pp. 151–161.
- [18] M. Leng and Y.-C. Wu, "Low-complexity Maximum-likelihood Estimator for Clock Synchronization of Wireless Sensor Nodes under Exponential Delays," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4860–4870, 2011.
- [19] B. R. Hamilton, X. Ma, Q. Zhao, and J. Xu, "ACES: Adaptive Clock Estimation and Synchronization using Kalman Filtering," in *Proc. of MobiCom 2008*, San Francisco, CA, 2008, pp. 152–162.
- [20] P. Lazik, N. Rajagopal, B. Sinopoli, and A. Rowe, "Ultrasonic Time Synchronization and Ranging on Smartphones," in *Proc. of RTAS 2015*, Seattle, WA, 2015, pp. 108–118.
- [21] S. Sur, T. Wei, and X. Zhang, "Autodirective Audio Capturing Through a Synchronized Smartphone Array," in *Proc. of MobiSys 2014*, Bretton Woods, NH, 2014, pp. 28–41.
- [22] C. Peng, G. Shen, and Y. Zhang, "BeepBeep: A High-Accuracy Acoustic-Based System for Ranging and Localization Using COTS Devices," *ACM Trans. Embed. Comput. Syst.*, vol. 11, no. 1, pp. 4:1–4:29, 2012.
- [23] Z. Li, W. Chen, C. Li, M. Li, X.-Y. Li, and Y. Liu, "FLIGHT: Clock Calibration and Context Recognition Using Fluorescent Lighting," *IEEE Trans. Mob. Comput.*, vol. 13, no. 7, pp. 1495–1508, 2014.
- [24] S. Schmid, G. Corbellini, S. Mangold, and T. R. Gross, "Continuous Synchronization for LED-to-LED Visible Light Communication Networks," in *Proc. of IWOW 2014*, Funchal, Portugal, 2014, pp. 45–49.
- [25] X. Guo, M. Mohammad, S. Saha, M. C. Chan, S. Gilbert, and D. Leong, "PSync: Visible light-based time synchronization for Internet of Things (IoT)," in *Proc. of INFOCOM 2016*, San Francisco, CA, 2016, pp. 1–9.
- [26] T. Akiyama, M. Nakamura, M. Sugimoto, and H. Hashizume, "SyncSync: Time-of-arrival Based Localization Method Using Light-synchronized Acoustic Waves for Smartphones," in *Proc. of IPIN 2015*, Banff, Canada, 2015, pp. 1–9.
- [27] L. Hou, S. Kagami, and K. Hashimoto, "An Advanced Algorithm for Illumination-based Synchronization of High-Speed Vision Sensors in Dynamic Scenes," in *Proc. of ICIRA 2010*, Shanghai, China, 2010, pp. 378–389.
- [28] R. Latimer, J. Holloway, A. Veeraraghavan, and S. Ashutosh, "Social-Sync: Sub-Frame Synchronization in a Smartphone Camera Network," in *Proc. of ECCV2014 Workshop on Light Fields for Computer Vision*, Zurich, Switzerland, 2014, pp. 561–575.
- [29] M. Sugimoto, H. Kumaki, T. Akiyama, and H. Hashizume, "High-Speed Optical Camera Communication Using an Optimally Modulated Signal," in *Proc. of ICASSP 2018*, Calgary, Canada, 2018, pp. 3739–3743.
- [30] IEEE Power Electronics Society, "IEEE Recommended Practices for Modulating Current in High-Brightness LEDs for Mitigating Health Risks to Viewers," <https://standards.ieee.org/ieee/1789/4479/> (retrieved: Oct 15th, 2022), 2015.