

# GNSS-based Time synchronisation of Stereo Images

Laure-Anne Gueguen<sup>1</sup>, Martin Wieser<sup>1</sup>, Gottfried Mandlbürger<sup>1</sup>

<sup>1</sup> Department of Geodesy and Geoinformation, TU Wien, Wiedner Hauptstraße 8/E120, 1040 Vienna, Austria  
(laure-anne.gueguen, martin.wieser, gottfried.mandlbuerger)@geo.tuwien.ac.at

**Keywords:** Stereo Photogrammetry, GNSS Time Synchronisation, ILX-LR1

## Abstract

Achieving precise synchronization in stereo photogrammetry is essential for accurately reconstructing dynamic scenes. A setup was developed that consists of two Sony ILX-LR1 cameras synchronously triggered using accurate timing provided by a GNSS receiver. For assessing the synchronization accuracy of the two cameras, we designed and implemented a board with LEDs providing visual reference. The evaluation indicates that the camera's latency between triggering and the actual image capture process is within 2.5 milliseconds (RMSE 562 us / max 2375 us / min -593 us). These minor latencies arise from internal camera processes that remain undocumented. Corrections could be applied based on the received signals from the camera to increase accuracy of the synchronization of cameras to be mostly better than 1 millisecond. The results validate the feasibility of precise synchronization for high-resolution, UAV-based stereo photogrammetry, ensuring minimal temporal discrepancies in image capture.

## 1. Introduction

Stereo photogrammetry for 3D reconstruction relies on the acquisition of images of the object from multiple viewpoints. In the specific case of a dynamic object, precise camera synchronization is crucial to avoid errors caused by temporal misalignment and to obtain an accurate 3D reconstruction. Even slight delays between camera exposures can introduce inconsistencies, particularly in scenarios, where fast moving objects or changing lighting conditions would lead to errors. Reconstruction of vegetation or trees in windy environments, reconstruction of the water surfaces, tracking of fluids, particles or objects in motion are among possible applications, which could benefit from a highly accurate camera synchronisation in photogrammetry.

To address this challenge, various methods have been developed. First, real-time hardware synchronisation methods provide the most accurate results. However, they are costly, require specific and complex hardware and lack flexibility, in the case of remote control for example. Such methods have been developed in (Holveck and Mathieu, 2004; Liu et al., 1997). Other alternatives rely on matching patterns between different images or video sequences, but they require suitable features, which are detectable and trackable through the sequence, and any matching error would cause an incorrect synchronisation. Another limitation is that they assume a constant timing offset between the cameras, which is not necessarily the case. They provide however an alternative when immediate trigger synchronisation is not possible and they minimise the requirements regarding data acquisition. (Rao et al., 2003; Sinha and Pollefeys, 2004).

In this context, we have developed a triggering setup for Sony's ILX-LR1 camera using the GNSS time as reference, which we would later like to integrate on a squad of Unmanned Aerial Vehicles (UAVs). We are aiming for an accuracy of the synchronization of different cameras of 2 to 3 ms. This paper is structured as follows: Section 2 introduces the equipment, Section 3 focuses on the timing challenge of our solution, Section 4 presents the test setup, Section 5 the associated results, and Section 6 a critical discussion.

### 1.1 Improving Long Baseline Stereo Photogrammetry

Specifically for designing a stereo photogrammetry setup for dynamic scenes mounted on moving platforms (UAV, car, pedestrian), synchronous image capture is of high importance. Most industrial cameras or specially designed photogrammetric cameras are normally used in high precise applications, as most consumer cameras lack clock stability and also proper electronic interfaces with low latency to precisely trigger images. The disadvantage of industrial cameras is mostly its high price or low image resolution compared to consumer cameras. Triggered by the emerging UAV sector, where lightweight but high-resolution cameras are in demand, Sony launched the ILX-LR1 camera in 2023. The aim is to build a bridge between consumer cameras and industrial cameras. In our setup, these Sony cameras have been chosen for their good balance in terms of resolution, weight, integrability on UAVs and cost. For UAVs with a payload capacity of less than 1kg, this camera fits well.

We chose the GNSS time as the basis for synchronous triggering of long-baseline stereo images because of its high accuracy. This entails the advantage to completely decouple the individual camera devices. This also saves us either a wired connection (which would not be suitable for UAVs anyways) or a communication protocol, which would require internet access.

## 2. Hardware

### 2.1 Camera

Sony's ILX-LR1 ([pro.sony/en\\_IL/products/installable-cameras/ilx-lr1](https://pro.sony/en_IL/products/installable-cameras/ilx-lr1)) is a lightweight 61MP full-frame camera. It is based on the Alpha series, but has a minimalist design. There is neither a display nor an internal battery. The weight of the camera is about 250 grams.

The ILX-LR1 features a power and control terminal which is used to power the camera and control image capture. For image capture, three signal lines are provided. Focus and trigger signals can be used to steer focus of the lens and to initiate an image capture, while the exposure line provides a signal related to the image exposure. When using the mechanical shutter, the

exposure signal will change its level once the front curtain is fully opened (see Figure 2). The mechanical shutter traversal time is about 3.5ms (stated by the camera's manufacturer instead of 4ms written in the manual<sup>1</sup>).

## 2.2 Camera Interface Board

We prototyped a simple electronic board, which allows to trigger the camera based on a precise GNSS time reference and log the exposure signal from the ILX-LR1.

The main electronic components of the device are:

**Sparkfun's GPS Breakout** ([www.sparkfun.com/sparkfun-gps-breakout-chip-antenna-sam-m10q-qwiic.html](http://www.sparkfun.com/sparkfun-gps-breakout-chip-antenna-sam-m10q-qwiic.html)), a breakout board for a u-blox SAM-M10Q GNSS receiver ([www.u-blox.com/en/product/sam-m10q-module](http://www.u-blox.com/en/product/sam-m10q-module)), and an integrated antenna receiver module. The Sparkfun board allows to straightforwardly communicate with the GNSS receiver and to access the Pulse-Per-Second (PPS) signal for reliable time synchronisation. The u-blox SAM-M10Q module (Figure 1) is a multi-constellation code signal GNSS receiver. It comes with an open-source software library (SparkFun Electronics, 2024) to deal with all the SAM-M10Q communication. The accuracy of the PPS signal is stated with 30ns (RMS) and to be within 60ns 99% of the time after the receiver's internal time is aligned.



Figure 1. Basic camera interface board using a u-blox SAM-M10Q (on the red board) to get the GNSS time frame.

**Arduino's NANO ESP32** board ([store.arduino.cc/en-at/products/nano-esp32](http://store.arduino.cc/en-at/products/nano-esp32)), a development board for Espressif's ESP32-S3 ([www.espressif.com/en/products/socs/esp32-s3](http://www.espressif.com/en/products/socs/esp32-s3)) module. The ESP32-S2 is a low-cost and low-power system-on-a-chip microcontroller with a dual-core RISC-V microprocessor including integrated Wi-Fi plus Bluetooth and a USB-C connector. The board serves as the main processing board that implements the time reference, communicates with the GNSS receiver, triggers the camera and receives the exposure signal from the camera.

The components used can be replaced by any GNSS receiver that provides a PPS signal. Any modern microcontroller that provides real-time hardware and timer interrupts can be used for the main board. We chose the GNSS board because of its low price and Sparkfun's extensive library.

## 3. Timing and Firmware

### 3.1 GNSS Time Reference

The UTC time frame provided by the GNSS receiver is used as the time reference of the system. This is implemented combining the PPS Signal and the received GNSS messages. The PPS signal is a precise and accurate time signal and its first signal edge is accurate to approximately 30 ns within the global time frame for which the receiver is configured. The time frame can be any realization, whereby UTC from the USNO (US Naval Observatory) is set by default. As the PPS signal is only a single hardware event, the absolute UTC timing reference can only be established along with UTC time received by the GNSS receiver's message.

In the current implementation, the PPS signal is connected to the main controller and triggers a hardware interrupt. Within the interrupt routine, the counter value of a hardware timer is set to the absolute microseconds of the full second in which the PPS occurred. The absolute value of the UTC's full second is derived from the last receiver's message plus 1 second. As seen in Figure 2, the specific receiver message (containing UTC time) used here is transmitted always after the PPS signal (here at 1Hz interval) and corresponds to the previous navigation epoch. The message contains additional flags that indicate the validity of the time measurement. The next PPS interrupt is only used to update the counter value of the timer if a valid GNSS time is specified.

A hardware interrupt is an event, which forces the microcontroller to execute a sub routine immediately with a low latency. A timer counter is a hardware register which increments its value (counter value) on a configured frequency (here the timer is configured to increment every microsecond). The counter value is a 64 bit register and thus can store absolute microseconds for about 30 k years, which is convenient as no byte overflow must be treated. A hardware timer can be used to trigger hardware interrupts on a defined counter value.

### 3.2 Camera Capture Timing

The described timer is used to trigger interrupts on counter values corresponding to absolute time (e.g., every 1 second but given in absolute microseconds). Within the interrupt routine, the trigger signal line is pulled to ground, which initiates the camera's capture procedure. Approximately 20 milliseconds later, the camera will pull the exposure line to ground, indicating that the first curtain of the mechanical shutter system is fully opened.

The exposure signal line is also connected to a hardware interrupt and the current timer counter value will be stored within the interrupt routine. The absolute times of the trigger signal and the exposure are both known and logged to analyse the timing differences between the two boards and cameras.

For reasons of readability, the term exposure signal in this document corresponds to the received start of the exposure signal.

We only use manual focus, and the camera requires the focus to be pulled to the ground even in manual focus mode before an image can be released. The signal line of the focus was simply set to ground for the entire duration of the session.

<sup>1</sup> [helpguide.sony.net/ilc/2390/v1/en/contents/231h\\_power\\_control\\_terminal.html](http://helpguide.sony.net/ilc/2390/v1/en/contents/231h_power_control_terminal.html)

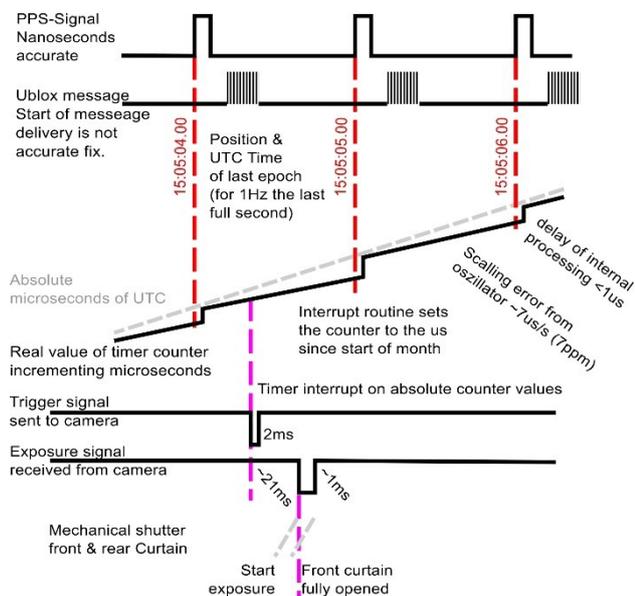


Figure 2. Timing Diagram of established GNSS timing reference and camera timings.

### 3.3 Internal Clock Error

For simplicity of the firmware, no calibration or correction of the internal clock source is done. That means that the internal timer counter (counting microseconds) is not exactly aligned to that from high precise GNSS time (as seen in Figure 2). The board's internal clock depends on the pulse of a physical clock source (mostly a quartz oscillator), whose inaccuracies result in a scaling of the clock. For the used board's oscillator, the specification states about 20ppm (parts-per-million), which is an error of 20 microseconds per second. That scaling error is temperature dependent and may vary between different pieces. But, as the timer counter value is updated every second on the PPS signal (on good GNSS reception), this error can be neglected as it is ten times smaller than the timing accuracy achieved between the cameras (as seen in Section 5.3). For our test environment, the measured clock scaling error is around 7 ppm for both boards and the temperature curves are also similar for both boards. As a result, the accuracy of the trigger signal based on the constructed time-frame between the boards is within 1 microsecond.

## 4. Test Setup

To investigate the timing accuracy and consistency between multiple cameras, two Sony ILX-LR1 cameras and two identically built camera interface boards (as described in Section 2.2) have been used. We executed several image capture sessions with different image intervals and durations. The tests have been carried out in an indoor environment in which the GNSS signals are provided via a GNSS repeater.

### 4.1 Camera Settings

The cameras were set to manual focus with an exposure time of 1/4000 s and f22 aperture. Only the mechanical shutter is used. Both cameras are configured identically by using a configuration file. For both cameras, we used Sony FE 35 mm F1.8 lenses. At the beginning of every image capture series, the lenses were focused on the centre of an LED panel, which served as visual reference.

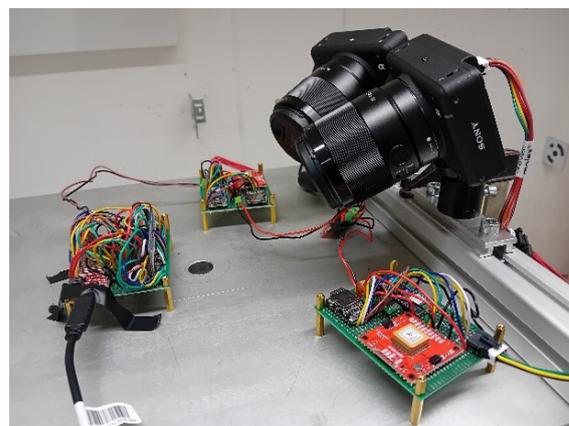


Figure 3. Setup with cameras, interface boards and the visual reference LED board.

### 4.2 Visual Reference

We not only investigated the accuracy and consistency of the cameras' exposure signal, but also used a visual reference for the image exposure time to detect inconsistencies or latencies that cannot be found by the pure electronic signals. Thus, we designed and implemented an LED panel (Figure 4) as a visual reference. The panel consist of 3x7 LEDs which are consecutively turned on and off for a configured time interval. The used on-off interval times are 400 us and 200 us, respectively, for each LED. The switching time between consecutive LEDs is less than 300 ns. The LED panel is connected to one of the camera interface boards and the LED sequence starts exactly 17ms after the trigger signal. The precision of the LED sequence is lower than 1us.



Figure 4. LED Board for visual reference of timing.

We applied a Python script to all images to extract all active LEDs. The processing steps include converting to grayscale, filtering out pixels with the lowest values (as the ISO and the exposure times had low values, making the images very dark), and finally extraction of the contours and their centres, using Open CV (OpenCV Team, 2024). Additionally, the timestamp from the EXIF (Camera & Imaging Products Association, 2024) metadata is extracted.

An example of the LEDs detection for a pair of images is shown in Figure 5. The first LED of each board counts as zero, which means the LEDs with Ids 1, 9 or 17 do not count in the sequence. In this example, LEDs 14-15-16 are blinking for the left image compared to only the 11-12 on the right image.

For each camera and image taken, the active LEDs are extracted and compared to each other and to the timing of the exposure signal. When the 400 us and 200 us LED intervals are used, along with 1/4000 s image exposure, there are mostly 1 to 3 active LEDs which are visible on the images.



Figure 5. LED detection for synchronised images: left camera (top) and right camera (bottom).

The viewing direction of the cameras was set so that the LEDs are approximately aligned horizontally at the image's middle height to minimise the timing difference due to the slow shutter traversal speed, as the shutter is traveling vertically with an approximate duration of 3.5 ms (see Section 2.1).

## 5. Analysis

### 5.1 Data Acquisition

In a first round of tests, several image acquisition sequences with different parameters were carried out to assess the feasibility of our idea and to check whether the prototype works. Table 1 summarises the tests.

Duration of acquisition	Exposure time	Frame rate	LED intervals
5min	1/4000 s	Every 1 s	400 us
7 min	1/4000 s	Every 1 s	400 us
20 min	1/4000 s	Every 10 s	400 us
1 h 45 min	1/4000 s	Every 1 s	400 us
18 h	1/4000 s	Every 10 s	400 us
22 h	1/4000 s	Every 10 s	400 us
15 h	1/4000 s	Every 10 s	200 us

Table 1. List of acquisition sequences; last row indicates the session using real-time correction (Section 5.4)

We started with very short tests and we then increased the total duration of the acquisition to test the stability of the triggering synchronicity through time, but also the repeatability and the robustness of the cameras' performances. We kept in mind that our final application is to perform drone surveys with several synchronised cameras of a river or a forest for example, therefore a representative configuration would be a total duration of 1 hour on average with images taken every second.

In particular, we executed three long-time acquisition sequences to reveal potential timing drifts or outliers, and to assess whether the cameras were able to sustain the command during a long period of time despite the potential temperature effects or camera internal delays for example.

### 5.2 Visual Reference compared to Exposure

First, the exposure signal indicating that the front curtain is fully open is analysed to determine whether it is tied to the physical process of image capture with a certain degree of accuracy. Therefore, the visual reference based on the LEDs is compared to the exposure signal received by the camera. The results are shown in the histogram in Figure 6. For the extracted time based on the LEDs, the centre time of all LEDs visible in an image is

used, where a single LED duration was 200 us for the session corresponding to the histogram.

The exposure signal for 11076 images (both cameras) is precise within a standard deviation of 185 us and an interquartile range of 92 us (the percentiles for 1% and 99% are -295 us and 204 us) to the visual reference's timing, and is reliable with just a few outliers (64 measurements  $> \pm 400$  us). The absolute time offset between these two time frames depends on the position of the image content due to the mechanical shutter traversal time, since only the time at which the front curtain is fully opened is provided. For our setup with an exposure time of 1/4000 s and an image content in the centre of the image, the pixels capturing the LEDs are already illuminated once the shutter is fully opened, as the traversal time is around 3.5 ms.

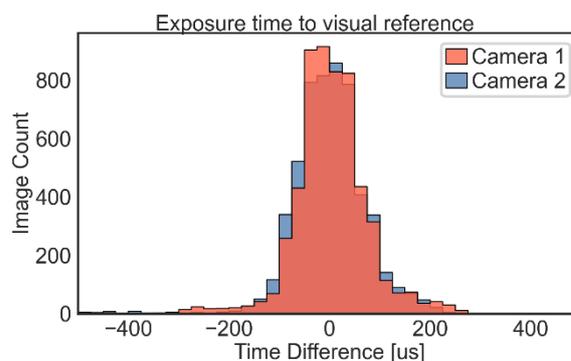


Figure 6. Histograms of the relative differences between the visual references compared to the exposure signals for both cameras (each 5538 images). The absolute time offsets of the visual reference and exposure signal have been reduced by their median.

The visual reference indicates that the exposure signal can be reliably used as the timing information of the physical image capture.

### 5.3 Camera Exposure Signal

Secondly, the consistency of the latency of the image exposure signal was analysed. The trigger signal, which is sent to the camera, is directly connected to a highly accurate global time. It is therefore important to know whether the offset between this trigger signal and the received exposure signal is constant over time.

In the end, for different cameras the received exposure signals (connected to the physical image recording) should be synchronized to a global time frame. In the used setup, the average nominal time difference between the trigger signal (sent to the camera) and the received exposure signal is about 20 ms.

As seen in Figure 7, there are variations of up 2400 us for both cameras in the time between the trigger and exposure signal (RMSE 562 us / max 2375 us / min -593 us). Some kind of periodical linear patterns are present but not at a fixed time or interval as both cameras have been turned on at the same time. The latency between the triggering and the received exposure signal is mainly positive and increases for both cameras compared to the nominal latency. However, the differences between the cameras on a global time frame are limited to around 2 ms.

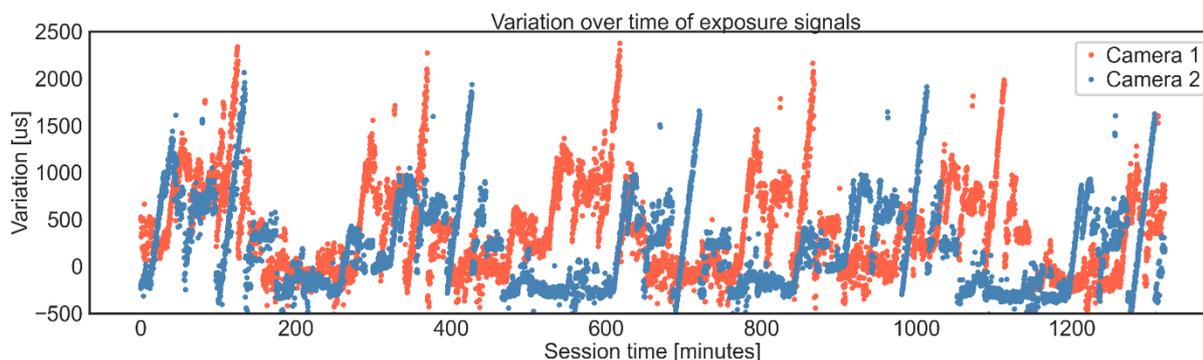


Figure 7. Variations of the exposure signal compared to the accurate trigger signal for both cameras (each 7918 images). Image interval: 10 s.

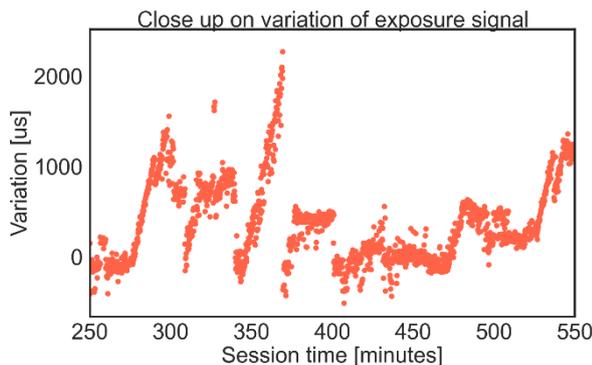


Figure 8. Close up on Camera 1 on the variations over of the exposure signal compared to the trigger signal, which is accurate to 2 us in relation to UTC. Image interval: 10 s

A closer look to the variations of Camera 1 in Figure 8 shows a relative slow change of these linear patterns, with some patterns introducing a constant offset while for others it seems that the exposure signal is drifting away for periods of more than half an hour. Abrupt changes and random fluctuations are present as well. The trigger signal sent to the camera is based on our established time frame and can be considered accurate to about 1-2 microseconds, which has been independently confirmed. As described in Section 4.2, the precision of the visual reference indicates that variations of the latency are not introduced by the interface board. For both cameras, the visual reference is precise to the exposure signal within 100 us, but the LED sequence is only started from one board. Therefore, it is most likely that the camera's firmware causes the documented variations. Some internal processes may require recurses, resulting in the varying latencies between the trigger signal and physical start of the image capture. Figure 9 and Figure 10 show the time differences between the received exposure signals of the two cameras in absolute UTC time. As already mentioned before, the real start of

the physical image capturing is disturbed by internal camera processes and thus this phenomenon is the main source of fluctuation in the time differences between the individual image exposures. All tests show the same behaviour. In particular, the Figures 7-10 present the results for the 22 hour-long test with 7918 images taken every 10 seconds. For these images, the RMSE is 710 us with an interquartile range of 990 us for the time differences in UTC between the exposure signals of the cameras (median -178 us / min -2344 us / max 1827 us.).

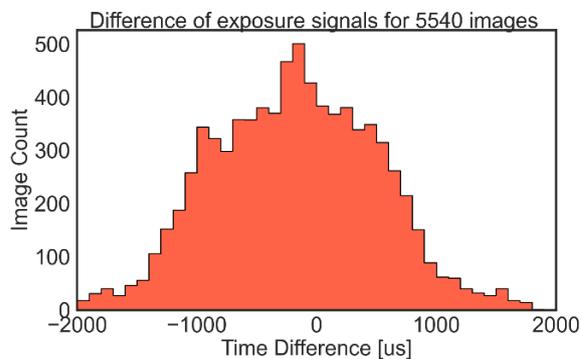


Figure 9. Histogram of the time differences of the received exposure signals in UTC between the two cameras.

#### 5.4 Correction of Camera Latencies

As seen and explained in the previous section, there are erratic behaviours in the camera's timing which cannot be explained. However, due to its relative linear dynamic in the sub-hour range, a basic correction is applied in real-time. The four last recorded time differences (with decreasing weights) between trigger and exposure signal are used to shift the start of the trigger signal in time so that the exposure signal will be aligned for different cameras in UTC. With this basic correction, the latency

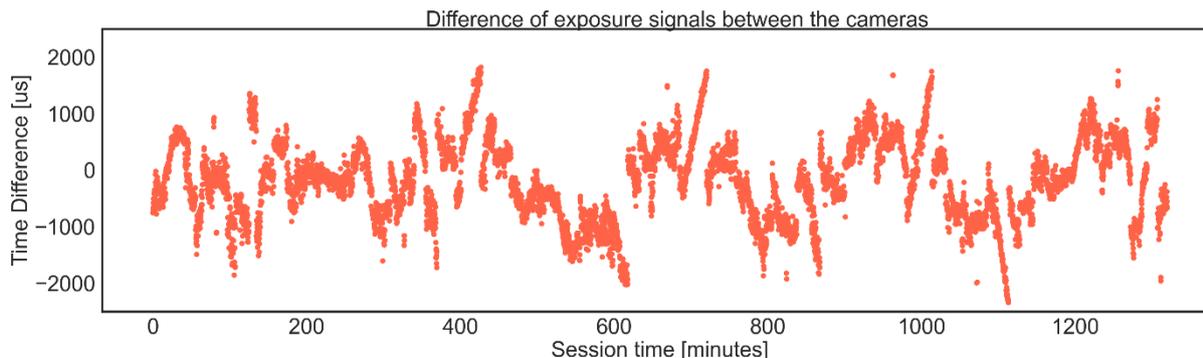


Figure 10. Time differences in UTC of the exposure signals between the two cameras

variations can be mostly compensated for the long-term patterns, even though some abrupt changes remain. Figure 11 shows the histogram of the differences between the exposure signals for both cameras in UTC with the real-time correction. The comparison between the histograms in Figure 9 and Figure 11, respectively, i.e. with and without correction, shows that this correction reduces significantly the time difference between the cameras' exposure signals. The accuracy of the time differences between the two cameras is then **470 us (RMSE)** with an **interquartile range of 329 us**, instead of 710 us and 990 us respectively for a similar dataset without correction, with a median of -1 us.

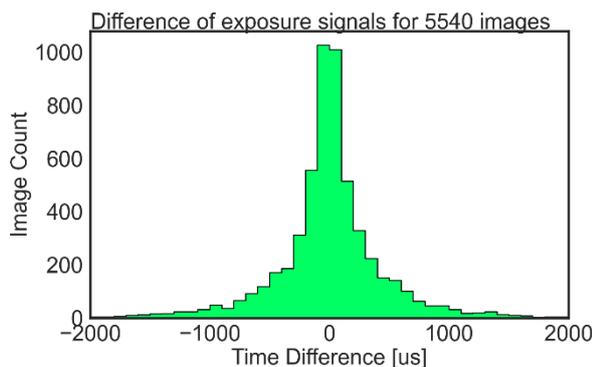


Figure 11. Histogram of the time differences of the received exposure signals in UTC between the two cameras with real-time latency correction.

## 6. Discussion

### 6.1 Internal Clock Drift

Considering outdoor settings, temperature changes, which affect all boards, will result in a similar timing error for all boards in the same direction and amplitude, which is negligible for short periods of 1 second after which the timer is again updated by the GNSS PPS signal. One problem remains, if one board has GNSS outages over longer periods while others are still updated by the PPS signal. The clock drift would then have significant impact on the absolute timing accuracy between the cameras. No investigations are undertaken to specify limits of acceptable GNSS outage periods, but a more decent firmware would solve longer GNSS outages by monitoring the internal clock drift and applying a correction term. Temperature dependent clock drifts are slow physical processes and thus can be modelled accurately.

### 6.2 Real-Time Exposure Timing Correction

Even though the correction of the latencies improves the overall accuracy, advanced modelling of the patterns during runtime could minimize outliers even more. On the other hand, for some cases it may just be enough to exclude those image pairs, which are not synchronized accurately enough. This can be done, by using the logged timings of the signals. Of course, the applicability of this strategy depends on the situation and the processes to be observed with the synchronized cameras.

### 6.3 Notes on Mechanical Shutter Traversal Time

As mentioned in Section 4.2, the time at which a specific pixel is illuminated depends on the traversal time of the mechanical shutter. Therefore, even though the cameras can be synchronized within sub milliseconds, depending on the pixel position of a 3D

object, the real synchronisation of that pixel for different cameras suffers from the shutter traversal time if they are not on the same vertical image position. This may not be a problem for static camera installations but for UAV-based stereo photogrammetry. Ideally, one can find mission parameters for the image's exterior run on different UAVs to minimize that synchronization error from the mechanical shutter's traversal time.

## 7. Conclusion

In stereo photogrammetry targeted at dynamic scenes, the synchronicity of the image capture by all cameras is crucial to obtain an accurate reconstruction. For UAV applications, there is also a growing need for lightweight and high-resolution cameras, to which Sony responded with the ILX-LR1 model. This model was chosen in the design of our multi-camera setup, where the cameras are triggered synchronously using accurate timing from a GNSS receiver as reference. A 3x7 LED board was used as visual reference to evaluate the performance of the synchronisation. Several tests were ran and the results are similar for all datasets. Overall, they showed that the latencies between the sent trigger signal and the received exposure signal are not constant over time but remain below 2.5 ms. These latencies are likely due to internal camera processes of which we have no knowledge. A simple correction algorithm is used in real-time to overcome exposure latencies by shifting the trigger signal in time to synchronize the received exposure signal of different cameras on a global time frame. The achieved accuracies for the two cameras then result in a RMSE of 470 us for the difference of the exposure time of the cameras. Finally, another advantage of the cameras used in this study is that the recorded timing of the signals can be used to eliminate image pairs where the latency of the exposure of one of the cameras is above a certain threshold.

## References

- Camera & Imaging Products Association, 2024. Exchangeable Image File Format for Digital Still Cameras: Exif Version 3.0. CIPA DC-008-2024-E.
- Holveck, B., Mathieu, H., 2004. Infrastructure of the GrImage Experimental Platform: The Video Acquisition Part. Technical Report No. 0301, INRIA Rhône-Alpes, Montbonnot-St-Martin, France.
- Liu, N., Wu, Y., Tan, X., Lai, G., 1997. Control system for several rotating mirror camera synchronization operation. Proc. SPIE, 2869, 695–703. <https://doi.org/10.1117/12.273473>.
- OpenCV Team, 2024. Contours: Getting Started. OpenCV Documentation. [https://docs.opencv.org/3.4/d4/d73/tutorial\\_py\\_contours\\_begin.html](https://docs.opencv.org/3.4/d4/d73/tutorial_py_contours_begin.html) (14 January 2025).
- Rao, C., Gritai, A., Shah, M., Syeda-Mahmood, T., 2003. View-invariant alignment and matching of video sequences. Proc. 9th IEEE Int. Conf. on Computer Vision, 939–945.
- Sinha, S.N., Pollefeys, M., 2004. Synchronization and calibration of camera networks from silhouettes. Proc. 17th Int. Conf. on Pattern Recognition ICPR 2004, Vol. 1, 116–119, IEEE Computer Society, Washington, D.C.
- SparkFun Electronics, 2024. A library which enables communication with u-blox GNSS modules. Version 3.1.7. [github.com/sparkfun/SparkFun\\_u-blox\\_GNSS\\_v3](https://github.com/sparkfun/SparkFun_u-blox_GNSS_v3) (28.10.2024).