Bachelor's Thesis

Czech Technical University in Prague

F3
Faculty of Electrical Engineering
Department of Cybernetics

Detection and Analysis of Periodic Movements Using an Event Camera

Jakub Kolář

Supervisor: Ing. Bc. Radim Špetlík
Study program: Open Informatics
Subfield: Artificial Intelligence and Computer Science
May 2024

BACHELOR'S THESIS ASSIGNMENT

I. Personal and study details

Student's name: **Kolá   Jakub**   Personal ID number: **507409**

Faculty / Institute: **Faculty of Electrical Engineering**

Department / Institute: **Department of Cybernetics**

Study program: **Open Informatics**

Specialisation: **Artificial Intelligence and Computer Science**

II. Bachelor's thesis details

Bachelor's thesis title in English:

Detection and Analysis of Periodic Movements Using an Event Camera

Bachelor's thesis title in Czech:

Detekce a analýza periodických pohyb   s využitím event kamery

Guidelines:

1. Perform a research on existing methods for measuring periodic motion using an event camera.
2. Describe the limits of data capture using the Prophesee EVK 4 event camera, for example the amount of contrast change the camera can report per second and the minimum time the camera can report a contrast change.
3. Create a set of recordings of periodic movements and changes from the event camera, at least for some of them with knowledge of the actual period. Aim for maximum variety of recording scenes, include for example rotating objects, shaking objects or a flickering light source.
4. Design and implement an algorithm to automatically detect periodic motion in event camera footage. Test the algorithm on the captured file and discuss its robustness.
5. Design and implement an algorithm to calculate the period of motion of an object captured by an event camera. Test the algorithm on a set of recordings taken under different conditions, for example, with a static camera and a moving object, or a static scene and a moving event camera, and determine its accuracy.

Bibliography / sources:

[1] Azevedo, George Oliveira de Araújo, Bruno José Torres Fernandes, Leandro Honorato de Souza Silva, Agostinho Freire, Rogério Pontes de Araújo, and Francisco Cruz. 2022. "Event-Based Angular Speed Measurement and Movement Monitoring" Sensors 22, no. 20: 7963. https://doi.org/10.3390/s22207963
[2] Zhao, Guanglei, Yiran Shen, Ning Chen, Pengfei Hu, Lei Liu, and Hongkai Wen. 2022. "High Speed Rotation Estimation with Dynamic Vision Sensors." ArXiv (Cornell University), September. https://doi.org/10.48550/arxiv.2209.02205.
[3] Natili, Francesco, Francesco Castellani, Davide Astolfi, and Matteo Becchetti. 2020. "Video-Tachometer Methodology for Wind Turbine Rotor Speed Measurement" Sensors 20, no. 24: 7314. https://doi.org/10.3390/s20247314

Name and workplace of bachelor's thesis supervisor:

Ing. Bc. Radim Špetlík   Visual Recognition Group   FEE

Name and workplace of second bachelor's thesis supervisor or consultant:

Date of bachelor's thesis assignment: **17.01.2024**   Deadline for bachelor thesis submission: **24.05.2024**

Assignment valid until: **21.09.2025**

_____   _____   _____
Ing. Bc. Radim Špetlík   prof. Dr. Ing. Jan Kybic   prof. Mgr. Petr Páta, Ph.D.
Supervisor's signature   Head of department's signature   Dean's signature

III. Assignment receipt

The student acknowledges that the bachelor's thesis is an individual work. The student must produce his thesis without the assistance of others, with the exception of provided consultations. Within the bachelor's thesis, the author must state the names of consultants and include a list of references.

_____._____
Date of assignment receipt

_____
Student's signature

Acknowledgements

I want to express my sincere gratitude to everyone who supported me throughout the completion of this bachelor's thesis. In particular, my brother for his assistance with many experiments, my parents for their everlasting support, and my supervisors, Ing. Bc. Radim Špetlík and prof. Ing. Jiří Matas, Ph.D., for their invaluable feedback, excellent guidance, countless hours of enriching discussions, and support throughout the research process. Finally, doc. Ing. Jan Fischer and Ing. Jan Fiala, thank you for providing us with the LED circuit board and control software used in one of the experiments.

Declaration

I declare that the presented work was developed independently and that I have listed all sources of information used within it in accordance with the methodical instructions for observing the ethical principles in the preparation of university theses.

Prague, date _____

_____
Signature

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne _____

_____
podpis autora práce

Abstract

This work presents a novel method for measuring properties of periodic phenomena (*e.g.*, rotation, flicker and vibration) using an event camera, a device asynchronously reporting brightness changes at independently operating pixels with high temporal resolution.

Our approach assumes that for a fast periodic phenomenon, a highly similar set of events is generated within a specific spatio-temporal window at a time difference corresponding to the phenomenon's period. The sets of similar events are detected by 3D spatio-temporal correlation in the event stream space.

The proposed method is evaluated on 12 sequences of periodic phenomena (*i.e.* flashing light and vibration) and periodic motion (*e.g.*, rotation) ranging from 3.2Hz (equivalent to 192 RPM) up to 2000Hz (equivalent to 120 000 RPM). The proposed method sets the new state-of-the-art in the measurement of periodic phenomena by achieving a mean relative error of 0.172%, which is significantly lower than the mean relative error of 32.252% achieved by the baseline method.

Keywords: Event camera, Frequency estimation, Periodic phenomena, Properties of periodic phenomena

Supervisor: Ing. Bc. Radim Špetlík
Karlovo náměstí 13 - building G
Prague 2, 121 35
Czech Republic

Abstrakt

Tato práce představuje novou metodu měření vlastností periodických jevů (např. rotace, blikání a vibrací) pomocí event kamery, což je zařízení asynchronně zaznamenávající změny jasu (eventy) na nezávisle pracujících pixelech s vysokým časovým rozlišením.

Náš přístup předpokládá, že pro rychlý periodický jev je v určitém časoprostorovém okně generován velmi podobný soubor eventů s časovým rozdílem odpovídajícím periodě jevu. Soubory podobných eventů jsou detekovány pomocí 3D časoprostorové korelace v prostoru eventů.

Navržená metoda je testována na 12 sekvencích periodických jevů (tj. blikající světlo a vibrace) a periodických pohybů (např. rotace) v rozsahu $3, 2 – 2000$Hz (což odpovídá 192 až 120 000 ot/min). Naše metoda dosahuje výrazně nižší průměrné relativní chyby $0, 172\%$, oproti $32, 252\%$ chybě jediné dostupné metody.

Klíčová slova: Event kamera, Odhad frekvence, Periodické jevy, Vlastnosti periodických jevů

Překlad názvu: Detekce a analýza periodických pohybů s využitím event kamery

Contents

Chapter 1

Introduction

Accurate measurement of periodic phenomena is crucial in diverse scientific and industrial fields. Precise quantification of rotational speed, for instance, finds applications in:

(i) monitoring of rotating components in machinery, *e.g.*, motor speed for performance evaluation and quality control [1],

(ii) aviation as drones require precise rotational speed monitoring for flight stability and maneuverability [2],

(iii) sports analysis, *e.g.*, ball tracking [3],

(iv) energy production as wind turbines rely on accurate rotational speed measurements to optimize energy production [4].

Traditional methods for measuring periodic phenomena often involve contact-based devices such as tachometers and rotary encoders. These approaches, while established, have inherent limitations: (i) physical contact with the target object can influence its movement and introduce measurement inaccuracies, (ii) contact methods may not be feasible in scenarios with delicate objects, confined spaces, or situations where interfering with the target significantly affects its movement.

Compared to contact measurement devices, laser tachometers offer highly accurate [5], less invasive measurements. However, reflective material (*e.g.*, a sticker) must be placed on the target to reflect the laser into the sensor while measuring. Under certain conditions, this limits the application of laser devices since it might not be convenient or even feasible to attach labels to particular surfaces or in confined spaces of the observed machinery. Another disadvantage is that the device operator must aim the laser precisely at the target, as missing the reflective material pass-through results in an inaccurate measurement.

Another crucial application of periodic phenomenon measurement lies in vibration monitoring. This technique, widely used in industrial settings, is vital in identifying imminent equipment failures [6]. By analyzing vibration patterns, engineers can assess the health of machinery and assets, preventing critical breakdowns and costly downtime. Predictive maintenance strategies leverage vibration data and analytics to detect anomalies before they escalate, enabling proactive repairs and minimizing production disruptions. However, conventional vibration monitoring technologies often rely on expensive equipment and require direct

contact with the target object [6], posing limitations similar to contact-based methods.

We propose a non-contact method that measures the properties of any periodic phenomena using an event camera. Our approach assumes that similar sets of pixel activations (events) will occur at periodic time intervals during the capture of the phenomenon and that these intervals correspond to the target frequency period. Put simply, we analyze 3D correlation responses to detect periodic similarities within the event data along the time axis. We evaluated our method on twelve sequences of four types of periodic phenomena: light flickering, object vibration, rotational speed, and periodic movement.

▌ Used abbreviations

- **GT** – Ground-truth

- **HD resolution** – High-definition resolution (1280×720px)

- **HDR** – High Dynamic Range

- **ROI** – Region of Interest

- **RPM** – Revolutions per Minute

Chapter 2

Event camera & Event-based vision

In recent years, event cameras and event-based vision have gained significant traction in machine vision for industrial automation, robotics, automotive, and other applications [7] as the ever-growing complexity of industrial environments demands innovative approaches to machine vision [8].

2.1 Differences between Frame-based and Event-based vision

The conventional approach, frame-based vision, is increasingly falling short in certain use cases [9], [10] and struggles with key challenges in industrial settings [11], such as capturing high-speed motion in dynamic scenes and handling scenes with extreme lighting variations.

An additional limitation of frame-based cameras is in their inherent nature: they capture entire images at fixed intervals (frame rate). Although this sufficiently replicates real-world action for human viewing, as individual frames at a frame rate higher than 24 frames per second appear to human eyes as continuous motion, the sensor constantly captures the whole scene, including static background elements, which generates a substantial amount of redundant data, while the information about the scene in between the frames is lost.

Event-based vision offers a compelling alternative. The asynchronous pixels of the sensor of an event camera continuously report changes in the scene brightness levels with microsecond temporal precision instead of colour information at regular intervals like a conventional camera. A comparison of frame-based and event-based camera outputs can be seen in Figure 2.1.

2.2 Prophesee EVK4 event camera & IMX636 sensor

We used the `Prophesee EVK4 HD` event camera in our experiments. The output resolution is 1280×720 pixels. The camera can capture up to 1066 million events per second [12] under light conditions ranging from 0.08 up to 100 000 lux [13] with microsecond time resolution (equivalent to 10 000 frames per second).

The behaviour of the `IMX636` camera sensor is adjustable with five biases [14], namely with two *contrast sensitivity threshold biases* (`bias_diff_on`, `bias_diff_off`),

two *bandwidth biases* (`bias_fo`, `bias_hpf`), and the *dead time bias* (`bias_refr`). The contrast sensitivity threshold biases regulate the contrast threshold, influencing the sensitivity of the sensor to changes in illumination. The `bias_diff_on` adjusts the `ON` contrast threshold, which is the factor by which the pixel must get brighter before a positive event is reported at that pixel, while the `bias_diff_off` determines the `OFF` contrast threshold, which is the factor by which the pixel must get darker before a negative event is reported at that pixel. Bandwidth biases control low-pass and high-pass filters, with `bias_fo` adjusting a low-pass filter to filter rapidly fluctuating light and `bias_hpf` adjusting the high-pass filter to filter slow illumination changes. Dead-time bias (`bias_refr`) regulates the refractory period of the pixel, determining the duration of non-responsiveness for each pixel after reporting an event.

In our experiments, we set the `bias_diff_on` contrast sensitivity threshold bias to 50 and the `bias_diff_off` bias to 30 due to a higher noise of events with positive polarities. Other biases remained at their default values.

The data acquired from the event camera are represented as a list of tuples (x, y, p, t), where x and y denote the spatial coordinates of the pixel that generated the event, t the timestamp in microseconds of event creation, and p is the polarity of the brightness change. A value of 0 indicates a decrease in brightness (negative event), while a value of 1 indicates an increase in brightness (positive event).

A fairly complicated readout process limits the accuracy of assigning a timestamp to an event in event cameras. When a pixel detects a contrast change, a readout request is emitted. The hardware identifies the row of the activated pixel, which is then scanned as a whole. Once the row is scanned, the hardware assigns a timestamp and polarity to each pixel, forming an event. The time between a real-world light change and the generation of an event is *pixel latency*, which is around 220 μs at 1000 lux and 310 μs at 10 lux [13].
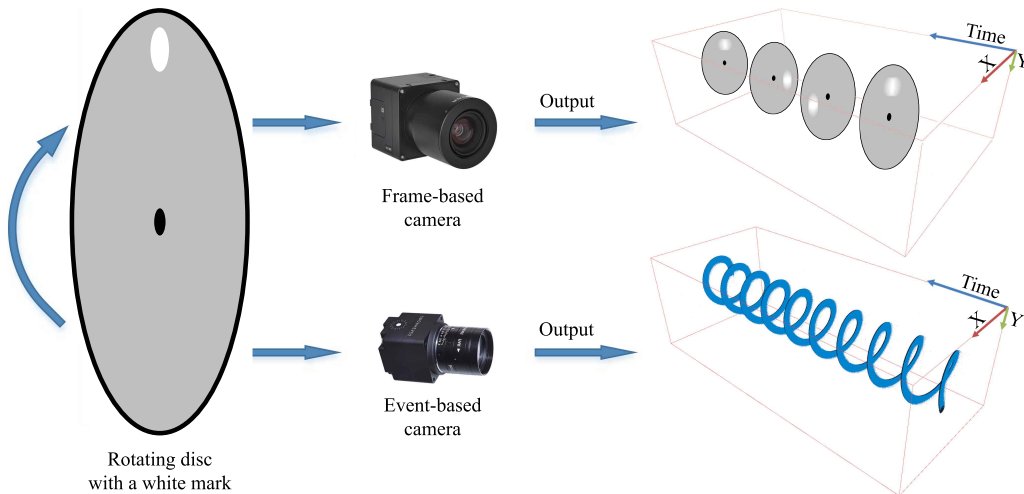

Figure 2.1: Rotating disc captured by frame-based (top) and event-based (bottom) cameras. In the same time span, a frame-based camera captures 4 RGB frames (top, time resolution 1/25 s), an event-based camera records $\approx 5 * 10^6$ events (bottom, time resolution ≈ 220 μs). Based on a figure in [15].

Chapter 3

Related work

3.1   Related work

In this section, we discuss existing approaches and technologies, mainly in the domain of rotation speed measurement, as periodic motion is arguably the most common periodic phenomenon. We categorise them into three groups: commercially available devices, camera-based methods using RGB cameras, and event-based methods utilising event cameras.

3.1.1   Commercially Available Rotation Speed Measuring Devices

Commercially available solutions for rotation speed measurement include contact and contactless methods. Contact-based devices include traditional mechanical tachometers that connect directly to the rotating shaft. These can introduce potential inaccuracies due to added mass and friction.

Electrostatic sensors achieve contactless measurement by detecting changes in the electromagnetic field generated by a rotating object equipped with a bearing and estimating the speed based on the frequency of these changes [16].

Optical encoder tachometers use a photoelectric sensor to detect light passing through a disc with alternating opaque and transparent segments, enabling speed estimation using the frequency of detected light changes [17].

Laser tachometers emit light that bounces off reflective stickers attached to the rotating target and measure the frequency of light detections back at the sensor [18].

3.1.2   Camera-based Rotation Speed Measurement Methods

Wang *et al.* [19] proposed a method using a low-cost camera and a simple marker on the target. Their method involves pre-processing sequential images, followed by similarity assessment and the application of the Chirp-Z transform to restructed signals. This approach achieves a relative error of $\pm 1\%$ in the speed range of 300 to 900 revolutions per minute (RPM).

An alternative approach presented in [20] involves calculating structural similarity and two-dimensional correlation between consecutive frames. The derived

5

parameters are used to reconstruct a continuous time-series signal, and the subsequent application of the Fast Fourier Transform allows for speed estimation with a maximum relative error of $\pm$ 1% up to 700 RPM.

To increase accuracy and broaden applicability, Natali *et al.* [21] utilised a sequence of coefficients derived from the correlation between a reference frame and subsequent frames. The short-time Fourier transform is then applied to compute the rotation speed. The algorithm was able to follow the measured speed up to the order of 1500 RPM with low error. As claimed by the authors, the major limiting factor is the frame rate of the acquisition rather than the data analysis method.

While offering the advantage of non-contact measurement, camera-based methods are often limited by the low frame rate of standard cameras. This can compromise accuracy, particularly for high-speed rotations, and limit the measurable rotation speed range. Additionally, in some cases, the requirement for observable landmarks or markers on the rotating object can restrict their applicability in specific scenarios where adding these is not feasible.

▪ 3.1.3 Event-based Rotation Speed Measurement Methods

Event cameras offer high temporal resolution, making them suitable for precise tracking of rapid rotational motion. Hylton *et al.* [22] introduced a method computing the optical flow of a moving object within an event stream, demonstrating its application in estimating the rotational speed of a disc with a black-and-white pattern. However, their algorithm lacked the sophistication for accurate high-speed rotation measurements.

Azevado *et al.* [23] introduced a simple method that allows for rotation speed measurements with a relative error lower than 0.2%, covering a speed range of up to 8500 revolutions per minute (RPM). The rotational speed is obtained by computing the elapsed time between spikes of distinct polarity events in a selected region caused by rotating patterns. However, the method struggles with targets that do not regularly produce an increased number of events with specific polarity (lacking high-contrast features) and is susceptible to noise.

The EV-Tach [24] method first eliminates events not produced by the rotating object by estimating the median distance from events to their centroid, eliminating events that exceed a specific threshold. Subsequently, it identifies rotating objects with centrosymmetric shapes and proceeds to track specific features, such as propeller blades. This method allows for speed measurements of up to 6000 RPM with a relative mean error as low as 0.3‰.

Chenyang *et al.* [25] proposed using laser-assisted illumination to help measure vibration frequencies to address the shortcomings of other vibration measurement event-based methods under unstable lighting conditions. At the same time, a mixture Gaussian distribution is proposed to fit the frequency distribution. Their method achieved successful vibration frequency measurement within a range of 20 to 300 Hz with minimal error (under 6.23‰) under ambient lighting of 0.01 up to 3800 lux.

Although event-based methods offer a temporal resolution for precise high-speed rotation tracking, their applicability can be limited in scenarios where the rotating

target lacks clear observable visual features or markers.

Given the lack of publicly available implementations for existing event-based methods for estimating the frequency of periodic phenomena, this work proposes two additional methods that serve as benchmarks for evaluating the performance of our proposed method. Those two methods are inspired by the concepts presented in the works mentioned above. Additionally, a closed-source method developed by the manufacturer of the utilised event camera is used as a third benchmark method. A detailed description of all methods is provided in the following chapter.

In conclusion, this section has explored various existing approaches for measuring rotation speed. While commercially available devices offer established solutions, camera-based and event-based methods provide non-contact alternatives with varying advantages and limitations depending on the specific application and target characteristics.

Chapter 4

Methods

When developing our method, we hypothesized that to accurately and reliably estimate the frequency of periodic phenomena, it is necessary to take into account not only the period of event generation in individual pixels but also the spatial pattern that the events create, which we explored in our previous work [26], on which we now follow up with an examination of spatio-temporal patterns in the event stream to improve the accuracy and range of applications further.

For the baseline method and our proposed method that uses 3D correlation, the spatial domain of captured event data is segmented into squares of equal size, referred to as *kernels*. We then analyse the event stream for each of these kernels separately. This allows us to present the most common or median result of per-kernel analysis to the user, resulting in a more robust approach. For computational efficiency, the user can define a two-dimensional area, *region of interest* (ROI), and analyse the event stream of only events with spatial coordinates within the kernels fully contained in ROI. Selecting a ROI results in reduced processing time, mainly when the periodic phenomenon of interest occupies a limited spatial area within the overall event stream.

4.1   Baseline method

To test our hypothesis that a method using individual pixels cannot accurately estimate the frequency of periodic phenomena, we developed a baseline method. As detailed in Section 2.2, event cameras produce a sequence of tuples (x, y, p, t), where x and y denote spatial coordinates, p represents the event polarity (positive – 1 or negative – 0), and t is the timestamp in microseconds.

To facilitate efficient event data storage, we used a four-dimensional sparse array, each dimension corresponding to a value in the tuples, to hold the event stream data. This enables data retrieval based on defined criteria, such as a specific spatial region, a particular polarity, or a selected time interval.

The sparse array is then used to extract all timestamps of events generated by a specific pixel and with a defined polarity. We refer to these timestamps as *PixPol-timestamps*.

The baseline method proceeds in three steps:

1. **Temporal analysis** – the delta times (Δ_t) between all consecutive PixPol-timestamps are calculated. Subsequently, the median or mode of these delta times is computed. We refer to the result as a *temporal analysis result* since it originates from data analysis along the temporal axis of the sparse array for a single pixel and polarity. The parameter of the method determines the choice between median and mode. This step generates two temporal analysis results per pixel (one for each polarity), resulting in $2 * kernelsize^2$ results per kernel.

2. **Spatial analysis** – temporal analysis results obtained in the previous step are aggregated. The median or mode of all the results within each kernel is computed. The second parameter of the method defines the function used. This step produces a single *spatial analysis result* per kernel, which estimates the period of the observed phenomenon in microseconds.

3. **Frequency prediction** – the final predicted frequency for each kernel is derived from the estimated period T in microseconds using the following equation:

$$\nu(T) = \frac{10^6}{T} \tag{4.1}$$

where $\nu(T)$ returns frequency in Hertz (Hz).

For plausible outcomes, the method assumes that most events are generated with the same periodicity as the period of the observed phenomenon.

The baseline method requires three parameters:

1. **Kernel size** defines the size of spatial areas where the frequency is estimated separately,

2. **temporal analysis function (median or mode)** is the method for summarising the delta times within a pixel and

3. **spatial analysis function (median or mode)** combines the temporal analysis results across all pixels in the kernel.

The evaluation of this method, including the exploration of various kernel sizes and combinations of analysis functions, is presented in Sec. 5.2.

▌ **4.2   Fast Fourier Transform method**

This section introduces a more robust baseline method for frequency analysis, using one-dimensional n-point Discrete Fourier Transform (DFT) with the efficient Fast Fourier Transform (FFT) algorithm [27].

To prepare the data for this analysis, we construct a binary time series based on PixPol-timestamps. In this series, a value of 1 is assigned to each position corresponding to a PixPol-timestamp, while all other positions remain at 0.

Put simply, this time series represents a sequence of spikes that occur at timestamps of event generation for a specific pixel and polarity. We refer to it as *event spike series.* To smooth out the spikes before applying the Fourier transform, we use the Hanning window. The Hanning window is defined as

$$w(n) = 0.5 - 0.5\cos(\frac{2\pi \cdot n}{M - 1}), \qquad 0 \leq n \leq M - 1 \qquad (4.2)$$

where $w(n)$ is the value of the window at the n-th sample, and M is the length of the window, in our case equal to 10^6.

Our objective is to estimate the frequency of event spikes within the time series, even when the events may not exhibit perfect periodicity. We do so by calculating the amplitude of the FFT frequency components and selecting the highest. In this way, we obtain the frequency at which events occur most frequently.

After calculating a 2D array of dominant frequencies for all pixels, we propose an optional post-processing step to eliminate potential outliers: applying a 2D median filter using a local window size of 3. Finally, we compute the most frequently estimated frequency.

This method, therefore, requires two parameters:

1. **Polarity** specifies events of which polarity to use in the creation of the event spike series and

2. **median filtering** defines whether median filtering should be applied.

The section 5.3 presents the quantitative results of this method with various configurations of parameters.

4.3   The Proposed method

This section introduces our proposed method, which extends beyond analysing the periodicity of events produced by individual pixels. The objective is to determine the period in which the events create a distinctive spatio-temporal pattern within the event data stream. We propose to compute a three-dimensional correlation of the event stream along the time axis with a selected subset of the event stream that we call a *template.*

By correlating the template (a chosen subset of the event stream) with the event stream, patterns occurring at different time offsets are found. For periodic phenomena, peaks in correlation responses are often found at regular intervals, as seen in Fig. 4.1. The time difference between these peaks represents the completion time of one period of the phenomenon. This analysis is performed for each kernel separately.

Given the inherent sparsity of the event data (for some experiments, events cover only 0.004‰ of the event stream space), we first transform it into a denser format. This is achieved by aggregating events with spatial coordinates within the kernel into a *3D spatio-temporal array* in which events are aggregated along the time axis within non-overlapping time windows of 100 microseconds. Each event is mapped to this array, with its value being 1 (positive polarity) or -1
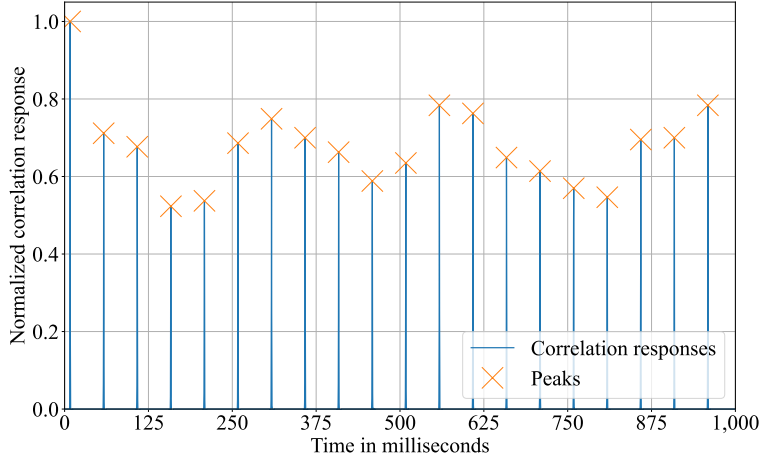
Figure 4.1: Normalized correlation responses of a selected 3D template with 500 ms of spatio-temporal event stream. Periodic peaks are highly distinctive, highly regular and indicate periodic phenomenon.
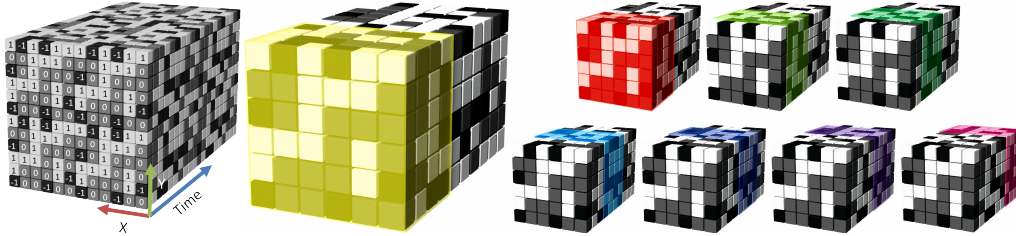

Figure 4.2: Dense 3D spatio-temporal array with positive events as white cubes, negative events as black cubes and no events as grey cubes (left), selected segment (middle) referred to as template is, put simply, used to compute 3D correlation with other same-sized segments (right) along the time axis of the 3D array.

(negative polarity). Other values in the array are set to 0. When multiple events occur within the same time window and pixel, only the polarity of the most recent event (with the highest timestamp) is retained. A visualisation of an example 3D spatio-temporal array is presented in Figure 4.2 (left).

The next step involves selecting a correlation template, a segment of the 3D array that will be used to compute the correlation over the entire array. The spatial size of the template is identical to the spatial size of the defined kernel. Its size along the time axis is a parameter that is adjustable for each kernel separately. An example of a template is illustrated in Fig. 4.2 (middle).

First, we experimented with a uniform template depth across all kernels. However, this method yielded inconsistent results, with the frequency being successfully measured in some kernels and failing in others. We hypothesize that this inconsistency is attributed to the inherent variability in event density across different regions of the event data stream.

To address this limitation, we adopted a more robust approach that adjusts the template depth within each kernel based on the event stream density. Specifically, we define a new parameter, *template event count*, which specifies the minimum
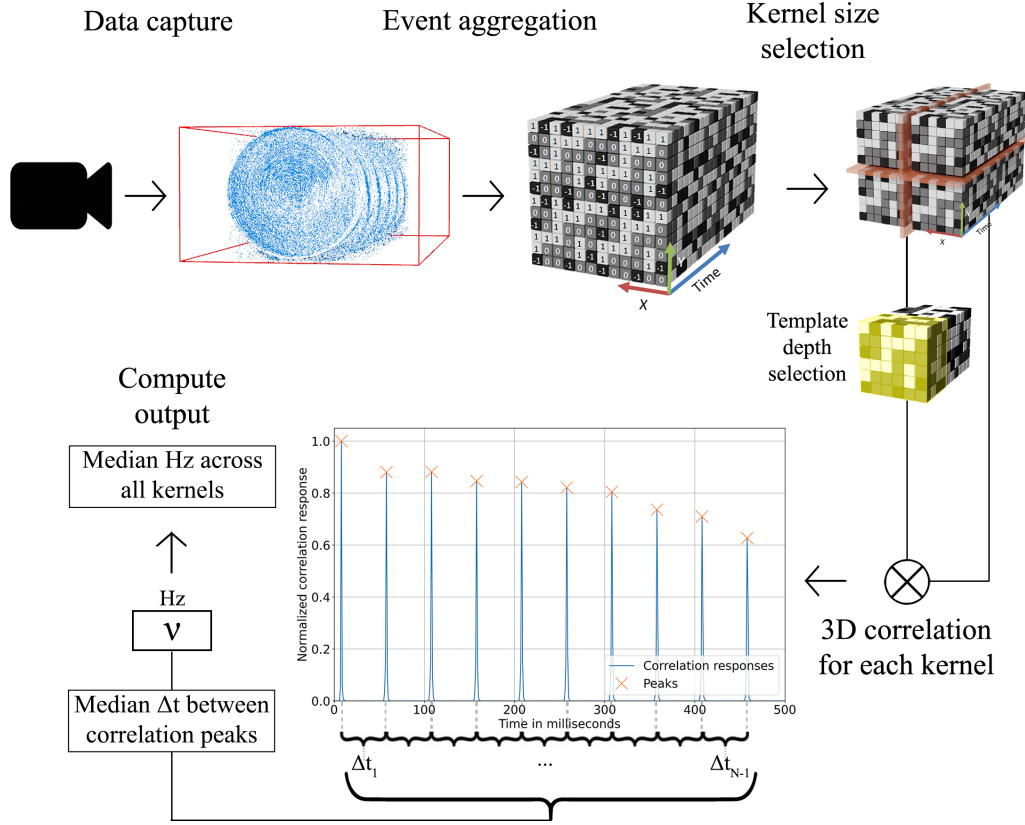
Figure 4.3: The proposed method: (i) data captured from an event camera is aggregated into a 3D array, (ii) a kernel size and a template depth are selected, (iii) 3D correlation of the template with the event stream is computed, (iv) a frequency is calculated from the average of time deltas measured between correlation peaks for each kernel, (v) the output frequency is computed as a median of measurements from all kernels.

number of events a template must contain. The template depth for each kernel is adjusted to ensure that it contains at least the minimum count of template events. This way, the template captures a sufficient representation of the local spatio-temporal pattern within each kernel.

The three-dimensional correlation between the template and the 3D array is then computed (see Fig. 4.2). The correlation responses exhibit periodic peaks, as expected, corresponding to the period of the observed phenomenon. The frequency is calculated by (4.1). An overview of the method is shown in Fig. 4.3.

To sum up, our proposed method requires two parameters:

1. **Kernel size** defines the size of square spatial areas where the frequency is estimated independently and

2. **template event count** specifies the minimum number of events a template must contain. The template depth for each kernel is dynamically adjusted to satisfy this requirement.

13

Figure 4.4–4.5 provides more insight in results of experiments 5.5.3 and 5.5.6. They contain:

- A visualisation of the event data over $\frac{3}{4}$ of one period of the phenomenon,

- a colour-coded grid of frequency estimations across all kernels and

- a sorted histogram of measured frequencies.

For all other experiments, see Fig. B.4–B.26.

For comparison with the results of other methods, a single numerical output is desired. By default, the method computes the median of estimated frequencies across all kernels.

(a) Event data   (b) Measured frequencies (Hz)   (c) Histogram of frequencies

Figure 4.4: Fronto-parallel velcro disc: (a) captured event data (positive events – white, negative events – blue), (b) grid of measured frequencies (blue – underestimated, green – correct, black – not enough events), (c) histogram of frequencies.

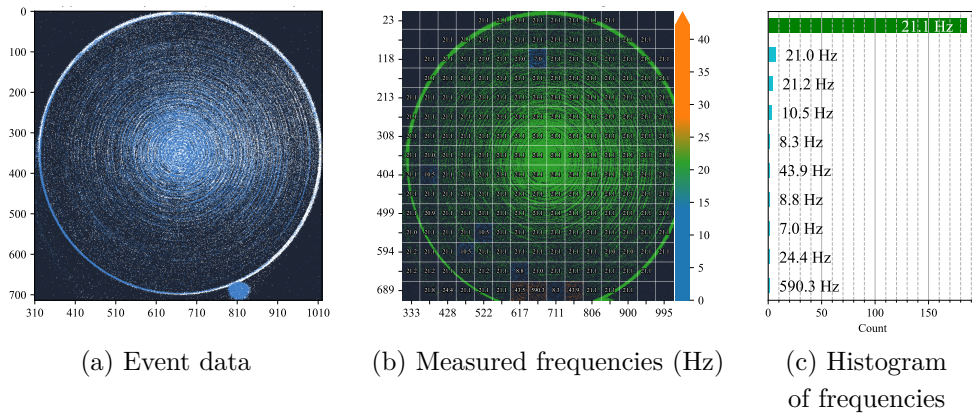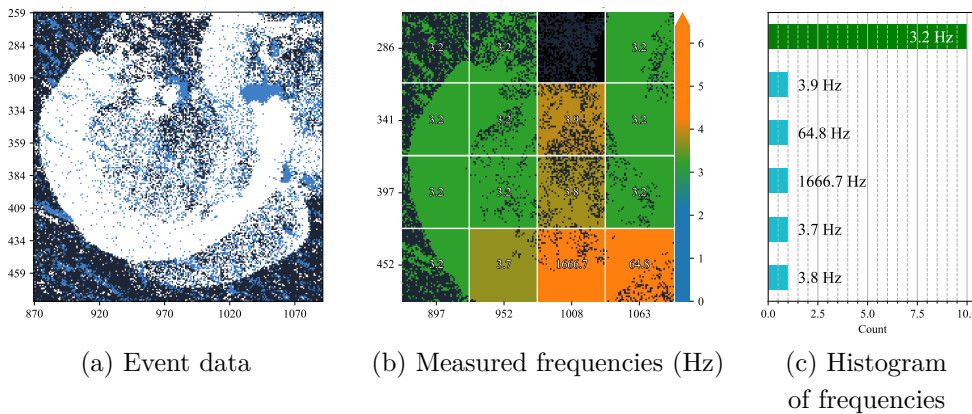(a) Event data   (b) Measured frequencies (Hz)   (c) Histogram of frequencies

Figure 4.5: Whirling Pholcus phalangioides: (a) captured event data (positive events – white, negative events – blue), (b) grid of measured frequencies (green – correct, orange – overestimated, black – not enough events), (c) histogram of frequencies.

Chapter 5

Experiments

In this chapter, we discuss the selection of parameters for methods presented in the previous chapter, followed by experiments measuring rotation speed, frequency of periodic light flashes, vibrations, and moving objects.

Due to the non-constant rotational speed of some targets, we restricted our analysis to one-second segments of the event stream. This selection was based on the assumption that the change in rotation speed within such a time frame would be negligible and not significantly impact the accuracy of our measurements.

To objectively compare the results of our proposed methods, we additionally use the publicly available closed-source vibration estimation method developed by Prophesee [28], manufacturer of the used event camera. This method provides a baseline comparison for our work, allowing us to assess the relative performance of our proposed approach.

When evaluating the accuracy of measurements, relative error and percent error are often used. The relative error quantifies the proportional discrepancy between the measured and actual values. It is calculated by dividing the absolute error by the actual value. This provides a dimensionless quantity independent of the units used in the measurement itself. The percent error expresses the relative error as a percentage. It is obtained by multiplying the relative error by 100%. When we refer to the "relative error" in our work, particularly when expressed as a percentage, we implicitly discuss the percent error.

5.1 Dataset

We used the `Uni-Trend UT372` laser tachometer [5] to capture ground truth (GT) rotation speed data. However, acquiring GT data this way was not feasible in all scenarios. In such cases, our EE3P method [26] was run multiple times with varying parameter configurations. This iterative approach aimed to filter potential measurement outliers and obtain a GT frequency estimate as close as possible to the actual value. The frequency was then verified by manually examining the event data stream. For each experiment, the method used to obtain the GT frequency, whether the laser tachometer or the EE3P method, will be explicitly stated.

Event data for three experiments in this work originate from the public dataset [29] created by Prophesee. For these experiments, we used the Event-to-Video method [30] developed by Prophesee to generate a reference photo of
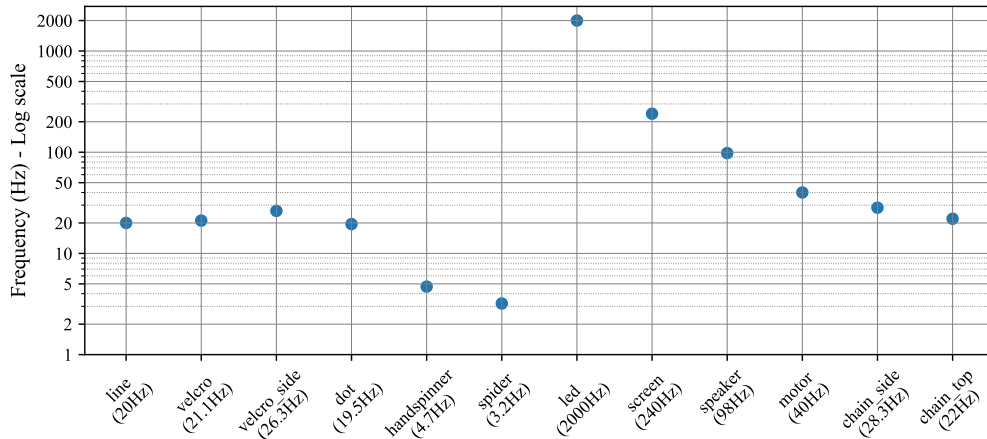
Figure 5.1: Ground-truth frequencies of experiments in order of appearance in this work.

the observed target, and the EE3P method [26] to estimate the GT frequency. We recorded all other sequences. The dataset contains experiments ranging from 3.2 Hz (equivalent to 192 RPM) up to 2000 Hz (equivalent to 120 000 RPM), see Fig. 5.1.

5.2 Selection of Baseline method parameters

To assess the impact of different parameter configurations on the performance of the baseline method (described in Sec. 4.1), we conducted an evaluation exploring all possible combinations of the following parameters:

1. **Kernel size** defines the size of the spatial areas where the frequency is estimated independently. The evaluated kernel sizes were 30×30px, 45×45px, 60×60px, and 75×75px, as we have not noticed significant improvements in further expanding or shrinking them,

2. **temporal analysis function (median or mode)** is the method for summarizing the delta times of events generation within a pixel and

3. **spatial analysis function (median or mode)** combines all temporal analysis results in the kernel.

The configuration achieving the best mean relative error (844%) across all experiments is the kernel size set to 75×75px and the median as both spatial and temporal analysis function. However, the applicability of this configuration was limited due to the size of one sequence (measuring phone screen refresh rate at an area of size 50×50px, see Sec. 5.6.2) being smaller than the proposed kernel size. Consequently, we opted to prioritise a kernel size applicable to all experiments.

Taking into account this constraint, the configuration that resulted in the third-best mean relative error (889%) while ensuring applicability to all data from our dataset was the kernel size of **45×45px**, and both the spatial and temporal

16

analysis functions were the **median**. This configuration successfully measured the dominant frequency in 4 of 12 experiments with a relative error below 5%.

The results for all parameter configurations of the method can be found in Table B.3.

5.3 Selection of Fast Fourier Transform Analysis method parameters

This section describes the selection of optimal parameters for the Fast Fourier Transform (FFT) analysis method (described in Sec. 4.2). We systematically evaluated all combinations of parameters of the method fixed across our dataset. These parameters are as follows.

1. **Polarity** specifies whether to analyze events with positive polarity or negative polarity when constructing the event spike series,

2. **median filtering** determines whether to apply a median filter to suppress outliers and improve the robustness of the method.

The configuration that used **negative events** and that **incorporated median filtering** produced superior results (mean relative error 60%) compared to other parameter configurations and the baseline method. This configuration successfully measured the dominant frequency in 6 of 12 experiments with a relative error below 0.5%. Refer to Tab. B.2 for the results for all combinations of parameters.

5.4 Selection of Three-Dimensional Correlation method parameters

This section discusses the selection of optimal parameters for our proposed method (described in Sec. 4.3). The parameters are:

1. **Kernel size** defines the size of the spatial area where the frequency is estimated independently. A larger kernel size can capture broader spatial patterns but might increase the computation time. The evaluated kernel sizes were 30×30px, 45×45px, 60×60px, and 75×75px, and

2. **template event count** specifies the minimum number of events a template must contain. A shallow template (low event count) might not capture the complete unique pattern of interest, leading to inaccurate frequency estimation. In contrast, a deep template (high event count) prolongs the computation time of 3D correlation.

To achieve a balance between capturing the relevant pattern and computational efficiency, we performed an analysis across all sequences in our dataset where we varied the minimum template event count and kernel sizes while measuring the following:

1. **Symmetric Mean Absolute Percentage Error (SMAPE)**: This metric measures the percentage difference between the observed and predicted values. SMAPE is calculated by finding the absolute difference between the actual and forecasted values and dividing it by the average of the absolute values of the actual and forecasted values:

$$\text{SMAPE}(y, \hat{y}) = \begin{cases} \frac{1}{n} \sum_{t=1}^{n} \frac{|y - \hat{y}|}{(|y| + |\hat{y}|)/2} & \text{if } |y| + |\hat{y}| > 0 \\ 1 & \text{otherwise} \end{cases} \tag{5.1}$$

where y represents the observed frequency, $\hat{y}$ represents the correct frequency and n represents the number of measurements we compare.

2. **Computation Time (Δt):** The time required to compute 3D correlation in all kernels using a specific parameter configuration. All calculations were performed on the same laptop using `Intel Core i7-9750H` CPU and `NVIDIA GeForce GTX 1660 Ti` GPU.

To evaluate the overall effectiveness of a parameter configuration, we introduce the *Performance-cost* measure

$$\eta(\text{SMAPE}, \Delta t; \lambda) = \text{SMAPE}^{\lambda} \cdot \Delta t \tag{5.2}$$

A lower value indicates better efficiency. We experimentally found that the ideal parameter λ is $\frac{3}{2}$ to prioritise configurations with lower SMAPE values at the cost of slightly longer computation times.

Figure B.1 illustrates the calculated SMAPE, the computation time, and the Performance-cost for various combinations of parameters for a single experiment.

We prioritise parameter configurations that perform well on computationally demanding experiments (those requiring more than 10 seconds to compute results for all kernels with a single-parameter configuration). For these experiments (5.5.1, 5.5.2, 5.5.3, 5.7.1), we calculated the sum of their Performance-cost values for each combination of parameters. The configuration with the overall minimum sum is considered the optimal choice. Figure B.2 presents the results. The best-performing parameters found are **kernel size 45×45px** and **minimal template event count of 1500 events**.

In the following sections (5.5-5.8), we will discuss experiments measuring rotation speed, frequency of periodic light flashes, vibrations, and moving objects.

■ 5.5 Measuring rotation speed

This section presents six experiments that showcase the versatility of our method for measuring the rotation speed of various objects under different conditions.

To capture ground truth (GT) rotation speed data, the `Uni-Trend UT372` laser tachometer [5] was used. The tachometer range is 10 to 99 999 RPM with a relative error of $\pm 0.04\%$. It is worth mentioning that the optical tachometer outputs only 3 to 5 samples per second, while our method produces a measurement for each detected period of the observed periodic phenomenon. We calculated the mean average of all the tachometer outputs captured in one second to obtain a single reference value for comparison with the results of all methods. However, in cases where acquiring GT data was not feasible, the EE3P method [26] to estimate the GT frequency, as described in Sec. 5.1.

■ 5.5.1 Felt disc with a high-contrast line

This experiment investigated the performance of our proposed frequency estimation methods in an environment with minimal noise. A power drill secured to a flat surface spun a white felt disc marked with a black line at a rotational speed of 1200 RPM, as shown in Figure 5.2a. The event data was captured simultaneously with readings from an optical tachometer, both positioned directly in front of the disc (fronto-parallel configuration).

Due to the nature of the experiment, characterised by minimal noise and near-accurate periodicity in event generation across pixels, we anticipated accurate frequency estimation from all methods used. The results confirmed this expectation.

The FFT analysis method produced accurate frequency estimates by analyzing event spike series of 98.6% pixels as shown in Figure B.3. Similarly, our method provided accurate results for all but one kernel, with only a minor deviation of 0.1 Hz observed in that specific case (see Figure B.4).

All four methods achieved relative errors below 0.05%, demonstrating their effectiveness in this controlled scenario. Prophesee's and our proposed method slightly outperformed baseline and FFT analysis methods in accuracy.

	Baseline	FFT Analysis	Prophesee Vibr. Est.	3D Correlation
Measured Frequency (Hz)	19.99	20.007	**20**	**20**
Relative Error (%)	0.05	0.034	**0**	**0**

Table 5.1: Measured frequency and relative error for the baseline, FFT, Prophesse Vibr Est and 3D correlation methods for experiment with a felt disc containing a high-contrast line (see Sec. 5.5.1).
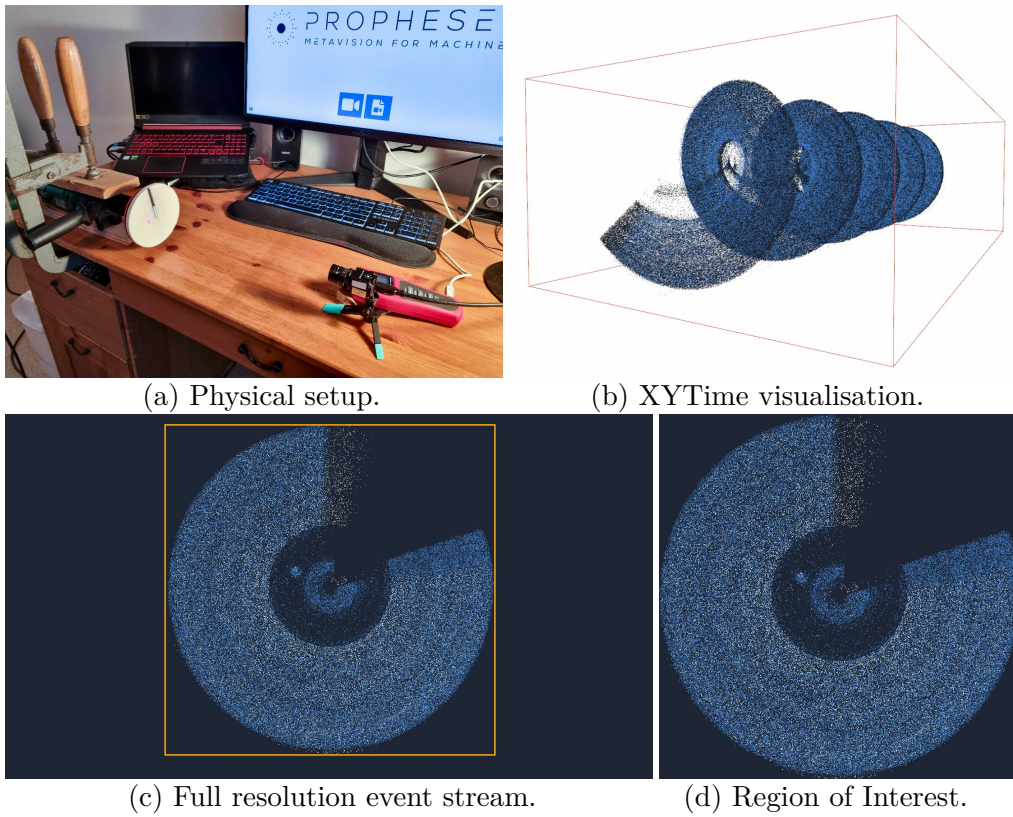
(a) Physical setup.

(b) XYTime visualisation.

(c) Full resolution event stream.

(d) Region of Interest.

Figure 5.2: Felt disc with a high-contrast line (see Sec. 5.5.1): (a) physical setup for shooting the experiment; (b) events from a 250-millisecond window visualised in spatio-temporal space; (c) aggregated events captured by the entire sensor (1280×720px) within a time window almost equal to the period of the observed phenomenon and highlighted Region of Interest (655×655px) shown in (d). Positive events are represented by white, and negative events are bright blue.

■ 5.5.2 Fronto-parallel velcro disc

This experiment evaluated the performance of frequency estimation methods on a more challenging scenario involving a disc covered in uniform velcro material, where pattern recognition is difficult even for the human eye.

The disc was rotated by a power drill at a speed of 1266 RPM with the event camera positioned directly in front of it and a laser tachometer measuring the rotation speed from one side of the disc using a reflective sticker attached to the edge of the disc (see Figure 5.3a).

This scenario demonstrates one of the advantages of event-based vision methods, as three of the four methods could detect and estimate the rotation frequency accurately. The inherent lack of periodic event generation across most pixels presented a challenge for the baseline method, as reflected in Table 5.2, which overestimated the frequency (by 14.576%). The FFT analysis method achieved the best estimate by analyzing event spike series generated by 6.1% of pixels that capture the edge of the disc (see Figure B.5).

Once again, both Prophesee's method and our proposed method successfully

estimated the frequency of the spinning velcro disc with 0% relative error, demonstrating their robustness in handling scenarios with minimal visual features (see Figure B.6).

	Baseline	FFT Analysis	Prophesee Vibr. Est.	3D Correlation
Measured Frequency (Hz)	24.176	21.005	**21.1**	**21.1**
Relative Error (%)	14.576	0.451	**0**	**0**

Table 5.2: Measured frequency and relative error for each method for experiment with fronto-parallel velcro disc (see Sec. 5.5.2).

(a) Physical setup.

(b) XYTime visualisation.

(c) Full resolution event stream.

(d) RoI.

Figure 5.3: Fronto-parallel velcro disc (see Sec. 5.5.2): (a) physical setup for shooting the experiment; (b) events from a 250-millisecond window visualised in spatio-temporal space; (c) aggregated events captured by the entire sensor (1280×720px) within a time window equal to the period of the observed phenomenon with highlighted Region of Interest (713×713px) shown in (d). Positive events are represented by white, and negative events are bright blue.

■ **5.5.3   Velcro disc with a non-frontal camera behind a glass sheet**

This subsection describes an experiment designed to further evaluate the robustness of frequency estimation methods under even more challenging conditions. A power drill rotated the velcro disc used in the previous experiment at a speed of 1578 RPM. The event camera captured data with a 45-degree camera angle through a sheet of glass (see Fig. 5.4a), simulating a possible industry application when the observed object is a rotating machine part and a transparent visor protects the camera and machine operators. The position of the laser tachometer remained the same. This experiment introduced additional complexity compared to the previous experiment due to the increased rotation speed, non-frontal camera position, and a transparent barrier between the camera and the target.

Interestingly, the baseline method estimated the frequency very well with only 0.39% relative error. The FFT analysis method achieved a relatively accurate frequency estimation with a 1.2% relative error. However, this frequency was only detected at 2% of all pixels, those that captured the edge of the disc (see Figure B.7). This led the method to output a frequency that was more commonly detected (by 11.9% pixels), this frequency being twice the actual frequency, likely due to the presence of centrosymmetric patterns on the velcro disc.

As in previous experiments, both Prophesee's method and our proposed method successfully estimated the frequency of the spinning velcro disc with 0% relative error, as shown in Table 5.3.

This experiment further highlights the robustness of our proposed method. Even with the non-aligned camera axis resulting in elliptical object trajectories on the image plane, the method produced accurate frequency measurements. Furthermore, the experiment confirms that capturing data through transparent materials does not compromise accuracy, as shown in Fig. B.8.

	Baseline	FFT Analysis	Prophesee Vibr. Est.	3D Correlation
Measured Frequency (Hz)	26.403	52.993	**26.3**	**26.3**
Relative Error (%)	0.39	101.495	**0**	**0**

Table 5.3: Measured frequency and relative error for each method for experiment with a velcro disc with a non-frontal camera behind a glass sheet (see Sec. 5.5.3).

(a) Physical setup. (b) XYTime visualisation.

(c) Target close-up. (d) Full resolution event stream. (e) Region of Interest.

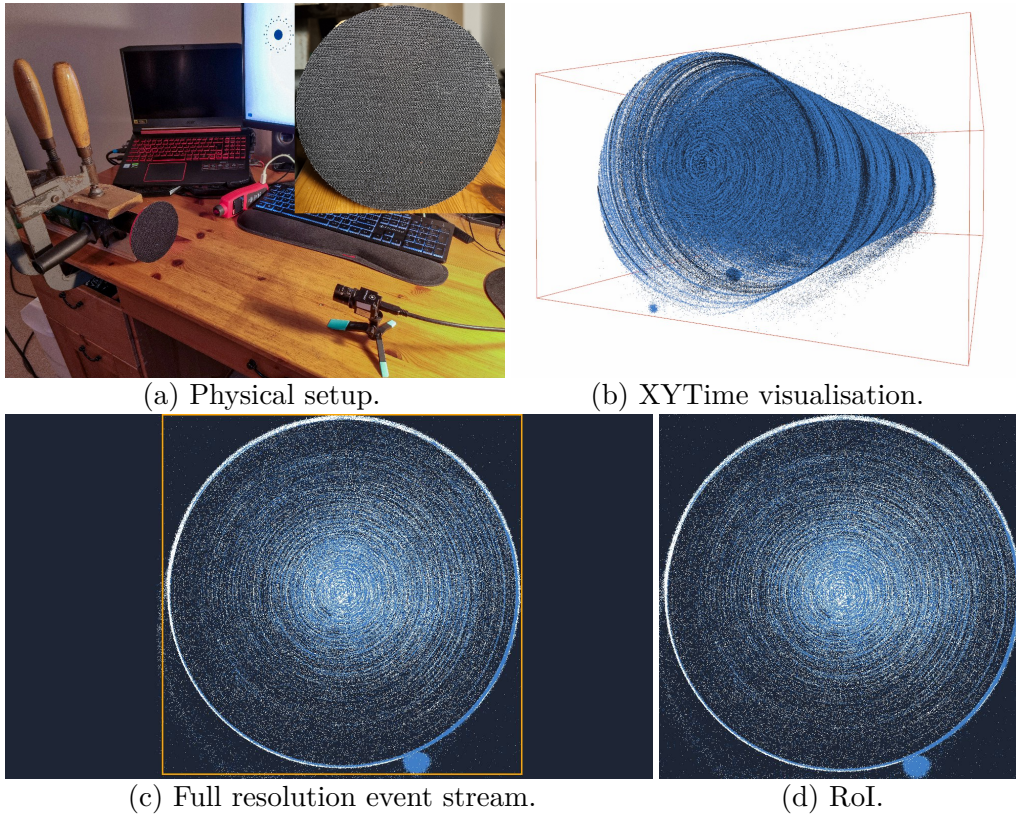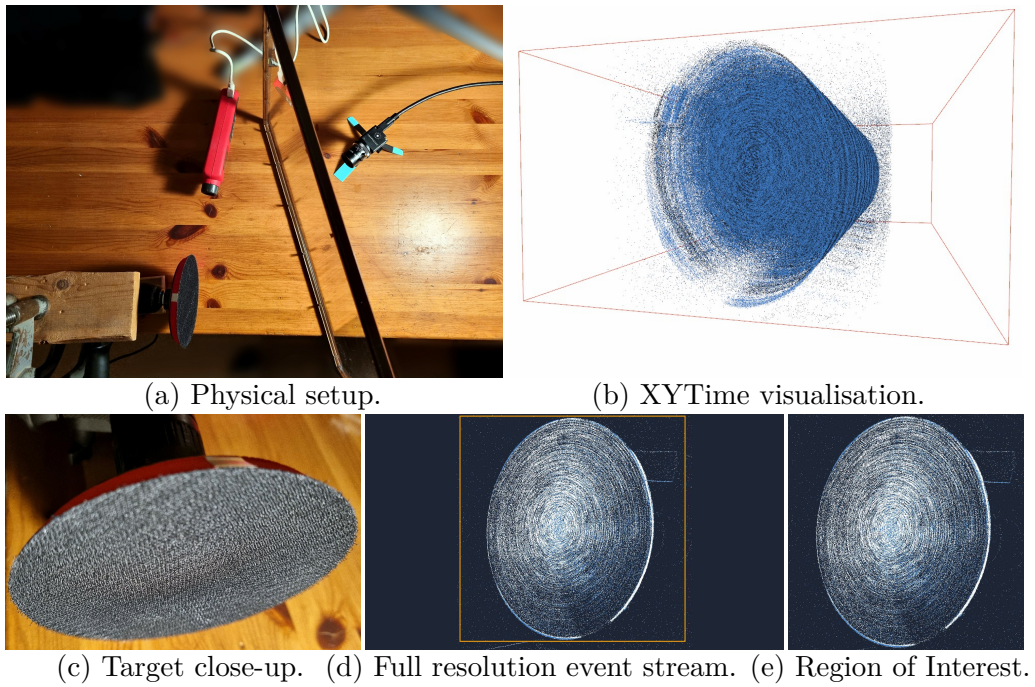**Figure 5.4:** Velcro disc with a non-frontal camera behind a glass sheet (see Sec. 5.5.3): (a) physical setup for shooting the experiment; (b) events from a 250-millisecond window visualised in spatio-temporal space; (c) a close-up photo of the target and reflective sticker on the edge of the disc; (d) aggregated events captured by the entire sensor (1280×720px) within a time window equal to the period of the observed phenomenon with highlighted Region of Interest (691×691px) shown in (e). Positive events are represented by white, and negative events are represented by bright blue.

5.5.4 High-contrast dot

This experiment evaluated the performance of frequency estimation methods on event data from Prophesee's public dataset [29]. The sequence captures an orbiting dot with a brightness lower than its surroundings. It was captured by an older model of their event camera with a lower resolution and an older sensor generation.

Since we did not acquire the event data ourselves, we lacked a photograph of the physical experimental setup. Therefore, the ground truth (GT) frequency was estimated using the EE3P method [26] and a reference photo was generated using the Prophesee Event to Video method [30] (see Figure 5.5a).

The baseline method did not produce any meaningful results, and the FFT analysis method estimated a frequency approximately three times higher than the GT frequency (see Figure B.9).

As in previous experiments, both Prophesee's and our proposed method (see Figure B.10) provided accurate estimates of the frequency of the observed periodic phenomenon, as shown in Table 5.4.

	Baseline	FFT Analysis	Prophesee Vibr. Est.	3D Correlation
Measured Frequency (Hz)	7462.687	58.984	**19.5**	**19.5**
Relative Error (%)	38170.188	202.484	**0**	**0**

Table 5.4: Measured frequency and relative error for each method for experiment with a high-contrast dot (see Sec. 5.5.4).

(a) Target visualisation.

(b) XYTime visualisation.

(c) Full resolution event stream.
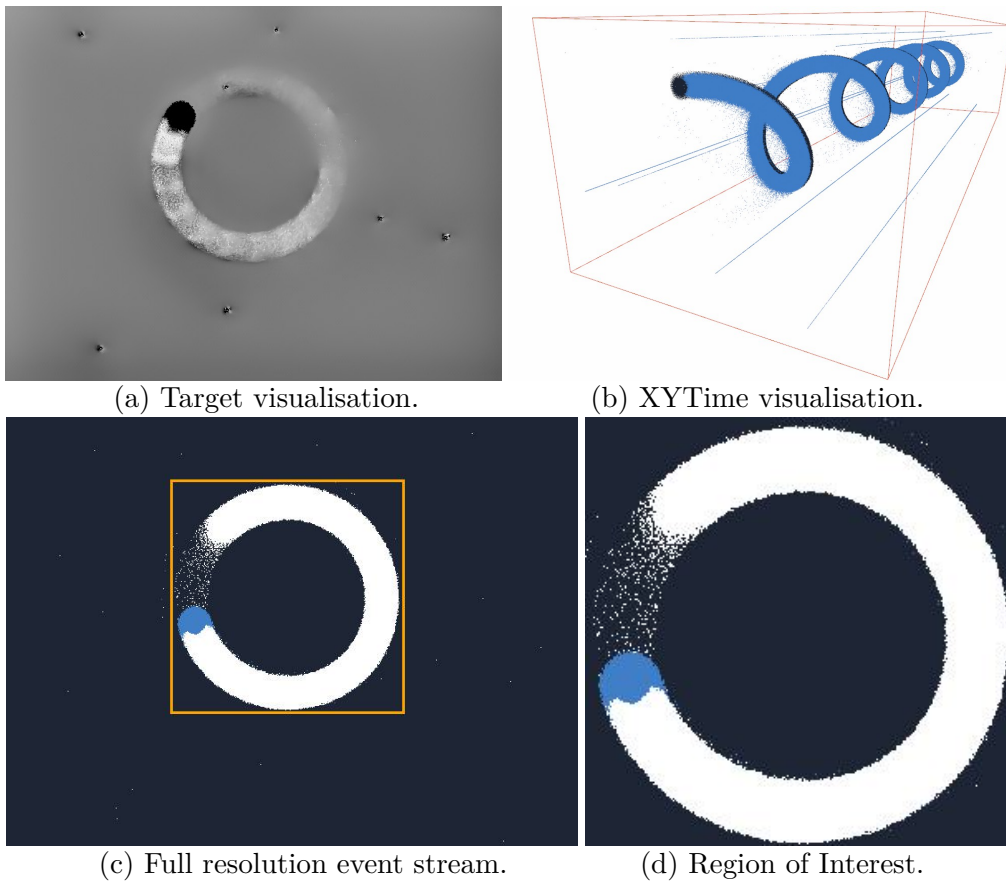
(d) Region of Interest.

Figure 5.5: High-contrast dot (see Sec. 5.5.4): (a) observed object visualised using the Prophesee's Event to Video method; (b) events from a 250-millisecond window visualised in spatio-temporal space; (c) aggregated events captured by the entire sensor (640×480px) within a time window almost equal to the period of the observed phenomenon with highlighted Region of Interest (263×263px) as shown in (d). Positive events are represented by white, and negative events are represented by bright blue.

■ 5.5.5 Hand fidget spinner

This experiment also utilized event data from Prophesee's public dataset [29], featuring a rotating three-blade hand fidget spinner. Similarly to the previous experiment, we used Prophesee's Event to Video method to generate a grayscale image of the target for reference (see Figure 5.6a). The ground truth (GT) frequency was estimated using the EE3P method [26]. The baseline method was unable to provide any meaningful results in this scenario. Despite the object being centrosymmetric, our proposed method and the FFT analysis method achieved good performance, as shown in Table 5.5.

The FFT analysis method successfully detected the correct frequency by analyzing event spike series generated primarily by pixels that capture the shadow of the fidget spinner (see Figure B.11). This observation highlights the method's potential to exploit even indirect event patterns for frequency estimation.

Interestingly, Prophesee's method produced a frequency approximately four times higher than the GT frequency. This discrepancy is particularly intriguing considering the *three*-blade design of the fidget spinner.

Our proposed method achieved the best overall performance, correctly estimating the frequency in 8 out of 10 valid kernels and exhibiting a deviation of only 0.1 Hz in the remaining two (see Figure B.12). This outcome emphasizes the robustness of our method in handling scenarios with centrosymmetric targets.

	Baseline	FFT Analysis	Prophesee Vibr. Est.	3D Correlation
Measured Frequency (Hz)	451.162	5	19.1	**4.7**
Relative Error (%)	9499.186	6.377	306.383	**0**

Table 5.5: Measured frequency and relative error for each method for experiment with a hand fidget spinner (see Sec. 5.5.5).

(a) Target visualisation.

(b) XYTime visualisation.

(c) Full resolution event stream.
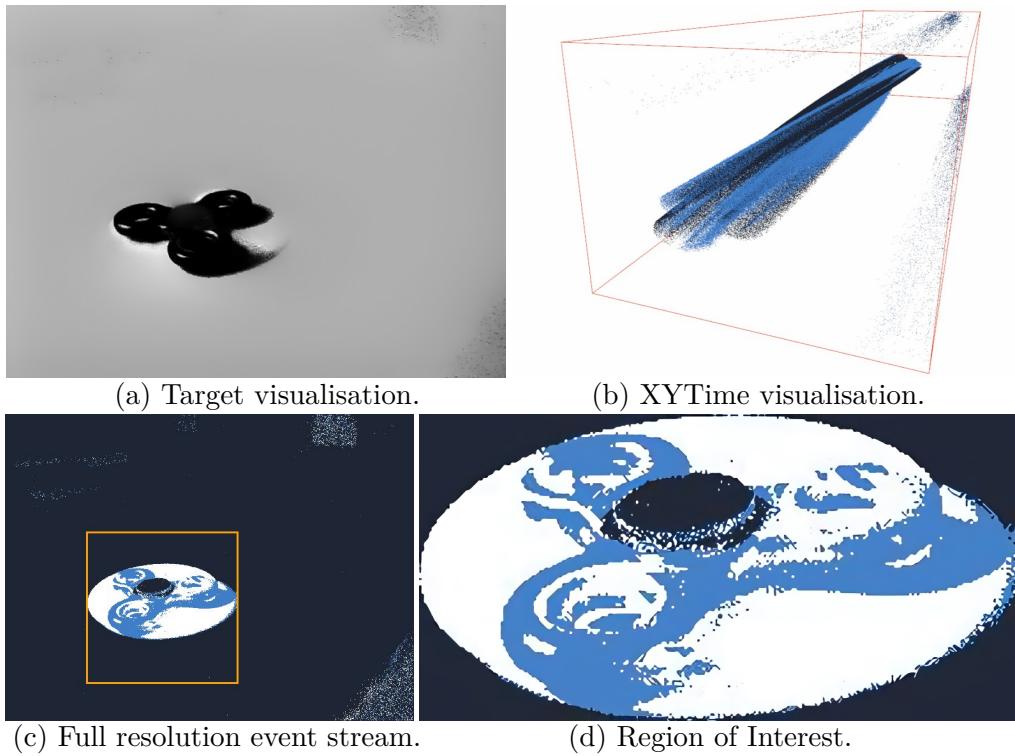

(d) Region of Interest.

Figure 5.6: Hand fidget spinner (see Sec. 5.5.5): (a) observed object visualised using the Prophesee's Event to Video method; (b) events from a 250-millisecond window visualised in spatio-temporal space; (c) aggregated events captured by the entire sensor (640×480px) within a time window equal to the period of the observed phenomenon with highlighted Region of Interest (239×239px) as shown in (d). Positive events are represented by white, and negative events are bright blue.

5.5.6 Whirling Pholcus phalangioides

This experiment investigated the whirling behaviour for evading predators used by the Pholcus phalangioides spider. This unique defense mechanism involves the spider rapidly rotating its body while its legs remain at the silk, creating a challenging yet intriguing scenario for measuring the rotation speed of the body of the spider. Additional details on this behaviour can be found in the work of R. R. Jackson *et al.* [31].

This experiment was characterized by significant noise, dim lighting conditions, and an unstable camera (see Figure 5.7a). Due to the light source being positioned on one side of the target, a second rotating pattern appeared, caused by the shadow of the spider (see Figure 5.7d,e).

While the baseline method failed to provide accurate results, the FFT Analysis method identified the correct frequency as the second most common. However, the frequency most frequently detected across all event spike series was four times higher than the reference frequency estimated using the EE3P method, leading to an inaccurate overall output (see Figure B.13).

Prophesee's method produced an estimate that deviated by 1 Hz from the GT frequency. Finally, our proposed method was the only one to achieve an accurate

result with 0% relative error, as shown in Figure B.14. This outcome highlights the effectiveness of our method in handling event streams with substantial noise.
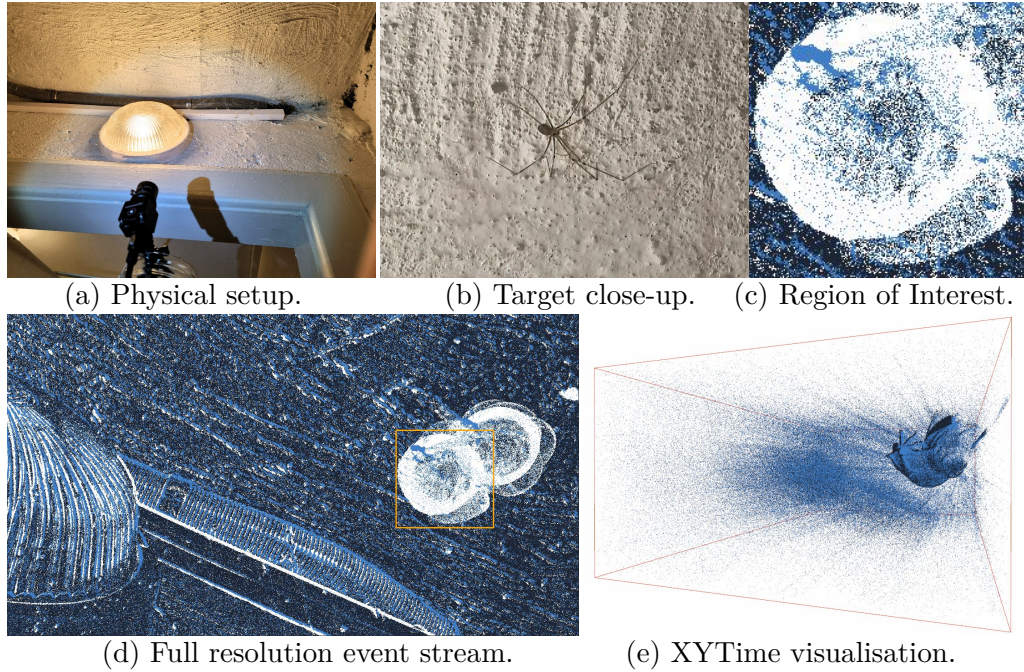

(a) Physical setup. (b) Target close-up. (c) Region of Interest.

(d) Full resolution event stream. (e) XYTime visualisation.

Figure 5.7: Whirling Pholcus phalangioides (see Sec. 5.5.6): (a) physical setup for shooting the experiment; (b) a close-up photo of the target; (d) aggregated events captured by the entire sensor (1280×720px) within a time window equal to the period of the observed phenomenon with highlighted Region of Interest (221×221px) as shown in (c); (e) events from a 250-millisecond window visualised in spatio-temporal space. Positive events are represented by white, and negative events are represented by bright blue.

	Baseline	FFT Analysis	Prophesee Vibr. Est.	3D Correlation
Measured Frequency (Hz)	35.135	12.993	4.2	**3.2**
Relative Error (%)	997.955	306.031	31.25	**0**

Table 5.6: Measured frequency and relative error for each method for experiment with a whirling Pholcus phalangioides (see Sec. 5.5.6).

27

5.6 Measuring frequency of periodic light flashes

This section presents the performance of frequency estimation methods in estimating high-frequency periodic light flashes.

5.6.1 Flashing LED

This experiment utilized a simple circuit with a light-emitting diode (see Figure 5.8a). A software oscilloscope on the Raspberry Pi Pico controller [32] allowed for precise control of the flashing frequency and duty cycle. The LED was set to flash at exactly 2000 Hz with a 50% duty cycle, challenging all methods to estimate a frequency substantially higher than those encountered in other experiments in this work or related publications.

The simplicity of this experiment lies in the near-synchronous event generation across almost all pixels. This scenario represents the only other experiment (after the fronto-parallel felt disc with a high-contrast line in Section 5.5.1) in which all methods achieved near-accurate results, as summarised in Table 5.7.

The FFT analysis method successfully estimated the frequency by analysing event spike series generated by 77.9% of all pixels. An additional 20% of pixels reported frequencies that were multiples of the ground truth (GT) frequency, further supporting the accuracy of the overall estimation (see Figure B.15).

The Prophesee's method produced a slight overestimation of 0.1 Hz, while our proposed method achieved a perfect estimate of the frequency of LED flashing, as shown in Figure B.16. This outcome further demonstrates the effectiveness of our method in handling scenarios with high-frequency periodic events.

	Baseline	FFT Analysis	Prophesee Vibr. Est.	3D Correlation
Measured Frequency (Hz)	1996.008	1999.412	2000.1	**2000**
Relative Error (%)	0.2	0.029	0.005	**0**

Table 5.7: Measured frequency and relative error for each method for experiment with a flashing LED (see Sec. 5.6.1).

28

(a) Physical setup.

(b) XYTime visualisation.

(c) Full resolution event stream.
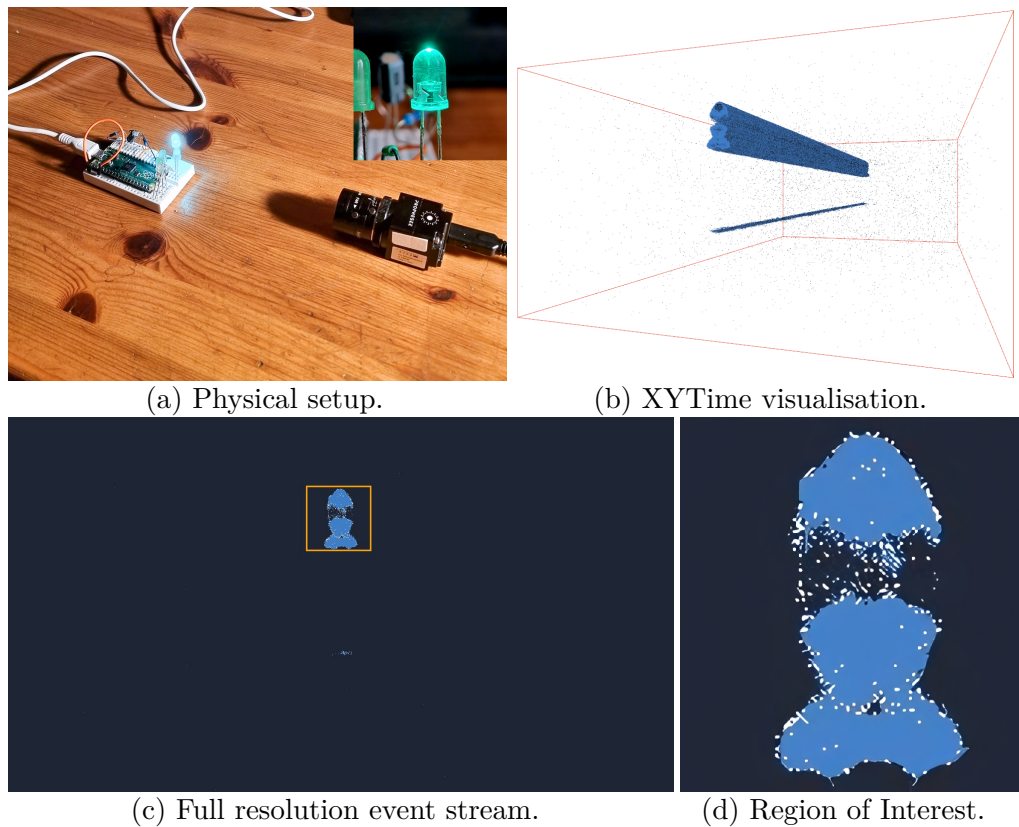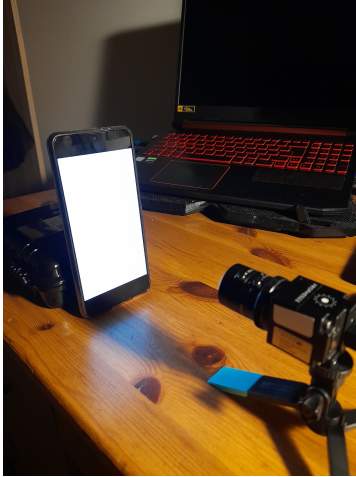

(d) Region of Interest.

Figure 5.8: Flashing LED (see Sec. 5.6.1): (a) physical setup for shooting the experiment; (b) events from a 250-millisecond window visualised in spatio-temporal space; (c) aggregated events captured by the entire sensor (1280×720px) within a time window equal to the period of the observed phenomenon with highlighted Region of Interest (126×126px) as shown in (d). Positive events are represented by white, and negative events are represented by bright blue.
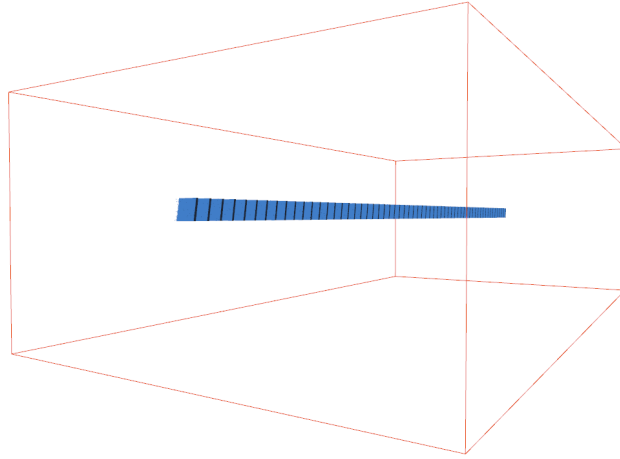
■ 5.6.2 Refreshing mobile phone screen

This experiment investigated the ability of frequency estimation methods to measure the flicker frequency of a mobile phone screen that displays a white colour at maximum brightness. Analysis of the event stream revealed periodic flashing patterns starting at the top and progressing towards the bottom of the screen (see Figure 5.9b). This behaviour is attributable to the Pulse Width Modulation (PWM) technique commonly used in displays to control backlight brightness. PWM involves rapid switching of the backlight on and off at a specific frequency. The flicker frequency was estimated using the EE3P method and confirmed by independent research on the display of this phone [33]. When capturing this experiment, we had to set a hardware-level region of interest to limit the observed area to avoid sensor saturation (see Figure 5.9c), hence the small RoI.

Although the baseline method failed to produce meaningful results, the FFT analysis method successfully accurately estimated the flicker frequency across all pixels (see Figure B.17).

The primary limitation of our proposed method lies in the process of aggregation of events within 100-microsecond time windows to generate the dense event array used by our method (as described in Sec. 4.3). This restricts the ability of our method to accurately measure periodic phenomena with periods not divisible by 100 microseconds. In this experiment, the exact period is $\frac{10^6}{240} = 4166,\bar{6}$ microseconds. Therefore, the closest estimate is 4200 microseconds, resulting in an estimated frequency of 238.1 Hz by our method, which translates to a relative error of 0.792% (see Figure B.18 and Table 5.8).

(a) Physical setup.   (b) XYTime visualisation.
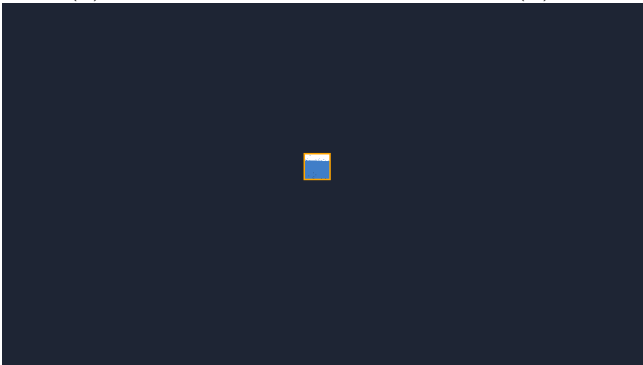
(c) Full resolution event stream.   (d) Region of Interest.

Figure 5.9: Refreshing mobile phone screen (see Sec. 5.6.2): (a) physical setup for shooting the experiment; (b) events from a 250-millisecond window visualised in spatio-temporal space; (c) aggregated events captured by the entire sensor (1280×720px) within a time window almost equal to the period of the observed phenomenon with highlighted Region of Interest (54×54px) as shown in (d). Positive events are represented by white, and negative events are represented by bright blue.

	Baseline	FFT Analysis	Prophesee Vibr. Est.	3D Correlation
Measured Frequency (Hz)	2770.083	240.02	**240**	238.1
Relative Error (%)	1054.201	0.008	**0**	0.792

Table 5.8: Measured frequency and relative error for each method for experiment with a refreshing mobile phone screen (see Sec. 5.6.2).

5.7 Measuring vibration frequency

This section presents the application of frequency estimation methods for measuring the vibrations of a speaker diaphragm and a vibrating motor. These scenarios show the potential of our proposed method for real-world industrial applications in vibration monitoring.

5.7.1 Speaker diaphragm

This experiment evaluated the performance of frequency estimation methods on the vibrating diaphragm of a speaker equipped with two large low-frequency drivers. To precisely control the emitted sound, we used an Android application [34] that allowed us to select and play a specific frequency.

We chose 98 Hz, corresponding to the musical note G_2 within the standard tuning system (meaning note A_4 is considered to be 440 Hz). The event camera captured the vibrating diaphragm of a speaker at approximately 30° angle. The most prominent features in the event stream originate from the vibrating edges of the diaphragm and the manufacturer's logo located in its centre (see Fig. 5.10).

The baseline method exhibited the most significant deviation from the reference frequency. The FFT analysis method achieved a relative error of 0.077% (see Figure B.19), and our proposed method achieved perfect accuracy (see Figure B.20). The Prophesee's method was the third most accurate, with a deviation of 0.2 Hz (see Table 5.9).

	Baseline	FFT Analysis	Prophesee Vibr. Est.	3D Correlation
Measured Frequency (Hz)	21.037	98.075	97.8	**98**
Relative Error (%)	78.534	0.077	0.204	**0**

Table 5.9: Measured frequency and relative error for each method for experiment with a speaker diaphragm (see Sec. 5.7.1).

(a) Physical setup.

(b) XYTime visualisation.

(c) Full resolution event stream.
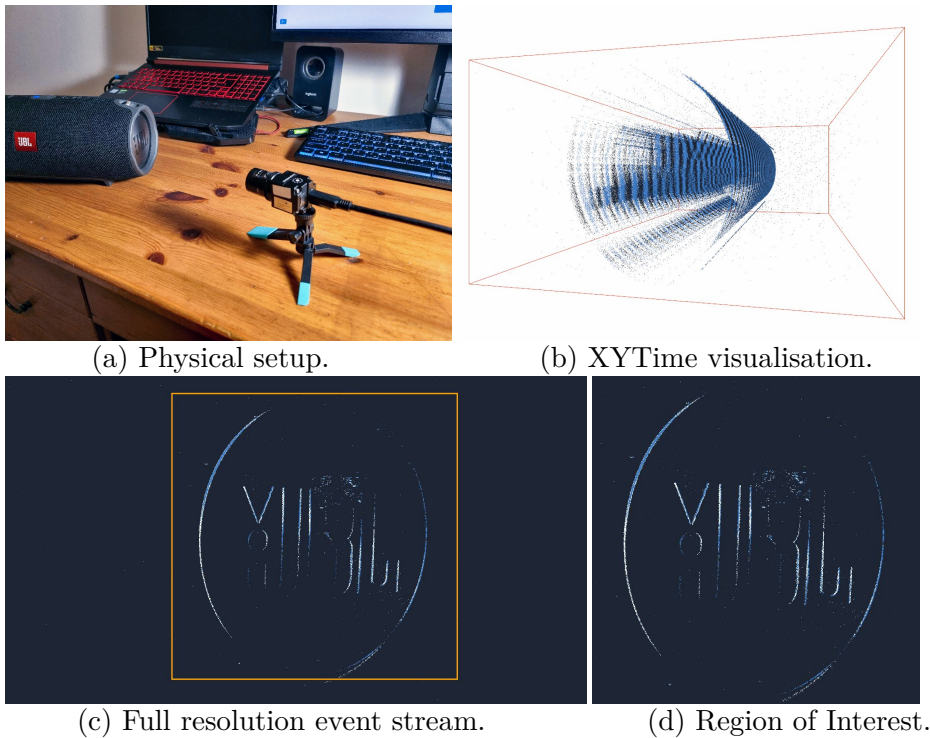

(d) Region of Interest.

Figure 5.10: Speaker diaphragm (see Sec. 5.7.1): (a) physical setup for shooting the experiment; (b) events from a 250-millisecond window visualised in spatio-temporal space; (c) aggregated events captured by the entire sensor (1280×720px) within a time window equal to the period of the observed phenomenon with highlighted Region of Interest (631×631px) as shown in (d). Positive events are represented by white, and negative events are represented by bright blue.

5.7.2 Vibrating motor

This experiment used event data from the public dataset [29] by Prophesee, where a motor vibrating at a constant frequency was captured. We used Prophesee's Event-to-Video method to generate a grayscale image of the target for visual reference (see Figure 5.11). The ground truth frequency was measured using the EE3P method [26].

The baseline method exhibited a significant deviation of 20% from the GT frequency. When evaluated across various event spike series, the FFT analysis method yielded accurate results in more than half of the cases, with the correct frequency identified. Furthermore, approximately 11% of the spike series led to the estimation of multiples (double and triple) of the GT frequency. The most reliable visual features this method used were the horizontal cooling fins of the motor, which vibrated periodically in a vertical direction (see Figure B.21).

Prophesee's and our proposed method (see Figure B.22) achieved perfect accuracy in estimating the vibration frequency. These findings highlight the ability of both methods to achieve accurate vibration frequency estimations even when the camera is positioned at different angles relative to the vibrating object, as demonstrated in previous and current experiments.

	Baseline	FFT Analysis	Prophesee Vibr. Est.	3D Correlation
Measured Frequency (Hz)	52.181	40.03	**40**	**40**
Relative Error (%)	30.453	0.075	**0**	**0**

Table 5.10: Measured frequency and relative error for each method for experiment with a vibrating motor (see Sec. 5.7.2).

(a) Physical setup.

(b) XYTime visualisation.

(c) Full resolution event stream.
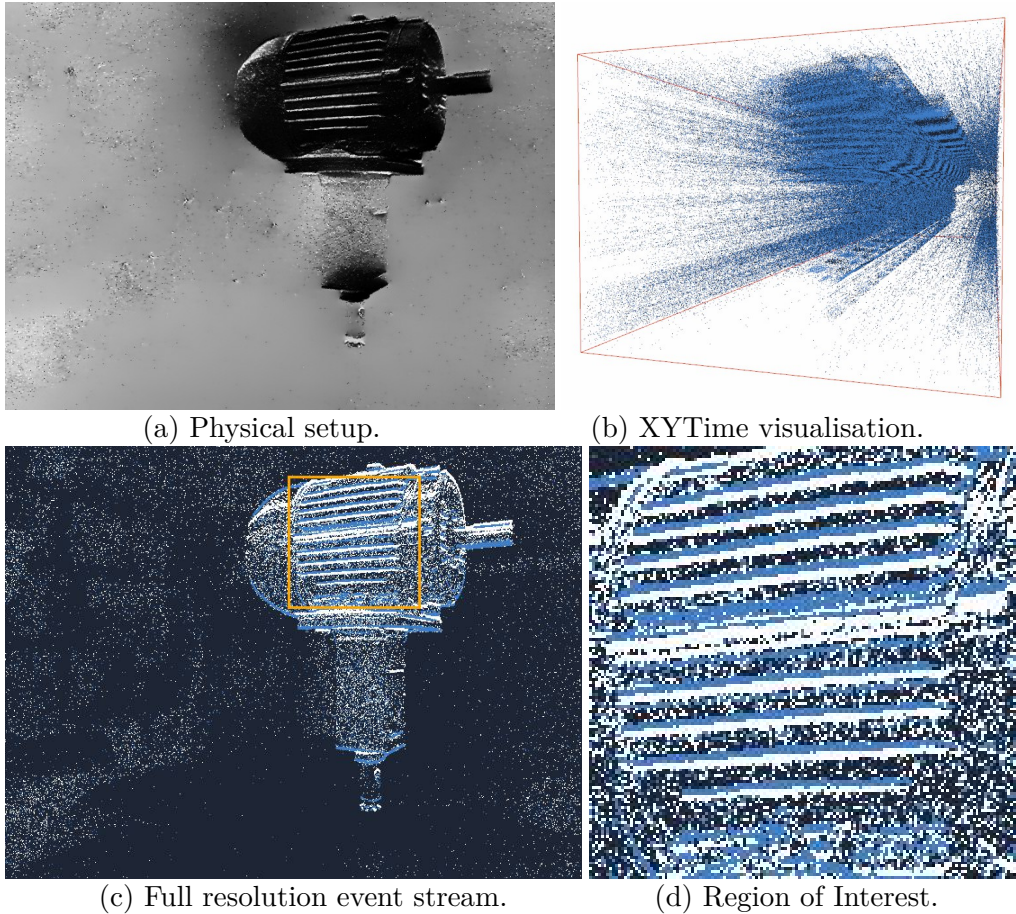

(d) Region of Interest.

Figure 5.11: Vibrating motor (see Sec. 5.7.2): (a) observed object visualised using the Prophesee's Event to Video method; (b) events from a 250-millisecond window visualised in spatio-temporal space; (c) aggregated events captured by the entire sensor (640×480px) within a time window equal to the period of the observed phenomenon with highlighted Region of Interest (148×148px) as shown in (d). Positive events are represented by white, and negative events are represented by bright blue.

5.8 Measuring movement frequency

This section presents the application of frequency estimation methods to measure the speed of moving objects passing by. We present two experiments with a moving bicycle chain observed from two distinct camera angles and estimate the frequency at which individual chain links pass by.

33

The experimental setup involved placing the bicycle wheels upright and spinning the pedals counterclockwise. The freewheel mechanism kept the wheels stationary while only the chain rotated, aiming to isolate the chain's movement for analysis.

■ 5.8.1 Bike chain from side view

This experiment captured a close-up view of the bicycle chain from a side angle, allowing for clear observation of individual chain links (see Figure 5.12). Due to technical difficulties, we had to set a sensor's hardware-level region of interest to lower the transfer speed required during the capture of this experiment.

Examining the results of the FFT Analysis method revealed that the areas reporting the correct frequency were mainly concentrated around the locations where the upper and lower portions of the chain links passed (see Figure B.23). This observation aligns with the expectation that event generation in these areas occurs almost exclusively once during individual chain link passages. However, these areas (32.5% of all pixels) constituted a minority compared to regions covering the middle sections of the chain links (53.7%), which produced incorrect frequency estimations of 1 Hz.

The Prophesee's method exhibited a deviation of approximately 18% from the GT frequency, marking the third experiment where this method failed to produce reliable results. The method did not identify a single dominant frequency, as analysis of each kernel yielded different frequency estimations. However, that does not necessarily compromise the accuracy of the method since the method, by default, outputs the median of all measured frequencies across kernels. As can be seen in Figure B.24, five out of 12 kernels reported frequencies ranging from 27.9 to 28.7 Hz, which falls within proximity of the reference frequency. As a result, our proposed method achieved the best performance among the four methods (see Table 5.11).

	Baseline	FFT Analysis	Prophesee Vibr. Est.	3D Correlation
Measured Frequency (Hz)	123.77	1.001	33.9	**28.7**
Relative Error (%)	331.254	96.513	18.118	**0**

Table 5.11: Measured frequency and relative error for each method for experiment with a bike chain from side view (see Sec. 5.8.1).

(a) Physical setup.

(b) XYTime visualisation.

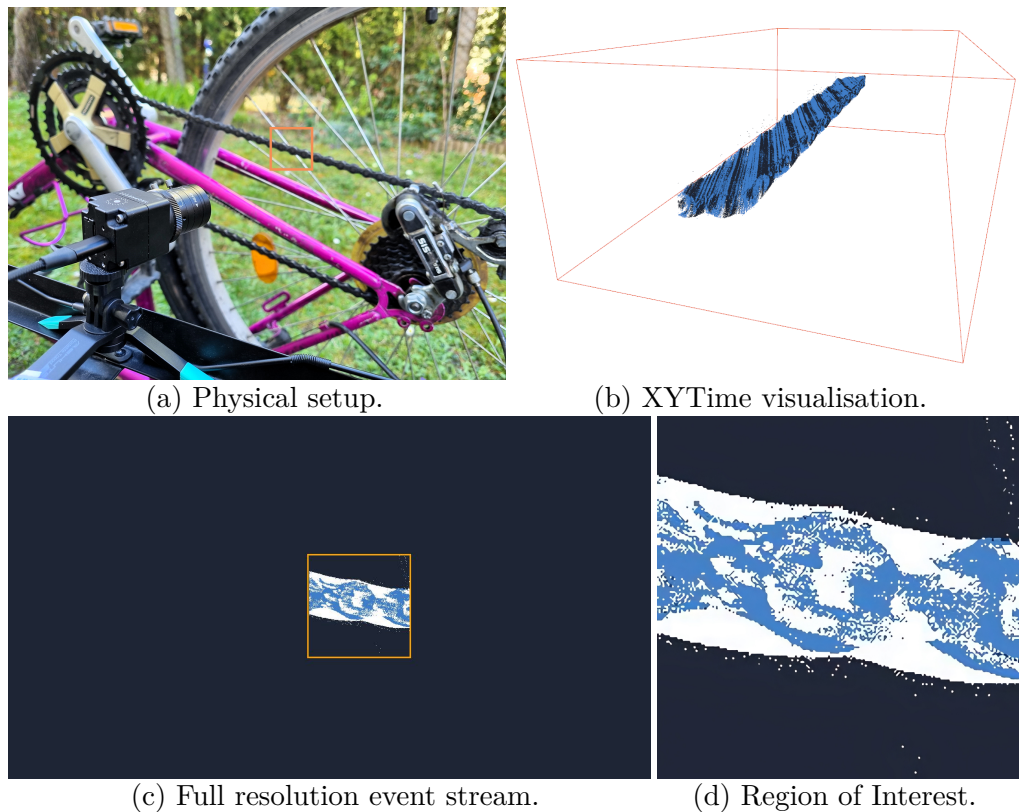(c) Full resolution event stream.

(d) Region of Interest.

Figure 5.12: Bike chain from side view (see Sec. 5.8.1): (a) physical setup for shooting the experiment; (b) events from a 250-millisecond window visualised in spatio-temporal space; (c) aggregated events captured by the entire sensor (1280×720px) within a time window equal to the period of the observed phenomenon with highlighted Region of Interest (204×204px) as shown in (d). Positive events are represented by white, and negative events are represented by bright blue.

■ 5.8.2 Bike chain from top view

The final experiment involved observing the bicycle chain from a top-down perspective (see Figure 5.13a), with the reference frequency estimated as 22 Hz using the EE3P method. Both rotating pedals were also captured but not analyzed and are visible in the event stream on the left and right side of the spatio-temporal visualisation in Figure 5.12b.

The FFT analysis method was the second most successful method in this experiment, reaching 6.4% relative error when analyzing event spike series generated by pixels capturing the components of the inner chain known as rollers (see Figure B.25).

Our method produced near-correct results between 22.8 Hz and 24.6 Hz in six out of nine valid kernels. However, their median produces the most significant deviation observed for this method in all experiments, with a relative error of approximately 1.3% (see Figure B.26 and Table 5.12). The baseline and Prophesee's methods exhibited significant deviations from the reference frequency.

	Baseline	FFT Analysis	Prophesee Vibr. Est.	3D Correlation
Measured Frequency (Hz)	51.645	21.993	30.8	**23.8**
Relative Error (%)	119.765	6.412	31.064	**1.277**

Table 5.12: Measured frequency and relative error for each method for experiment with a bike chain from top view (see Sec. 5.8.2).

(a) Physical setup. (b) XYTime visualisation.

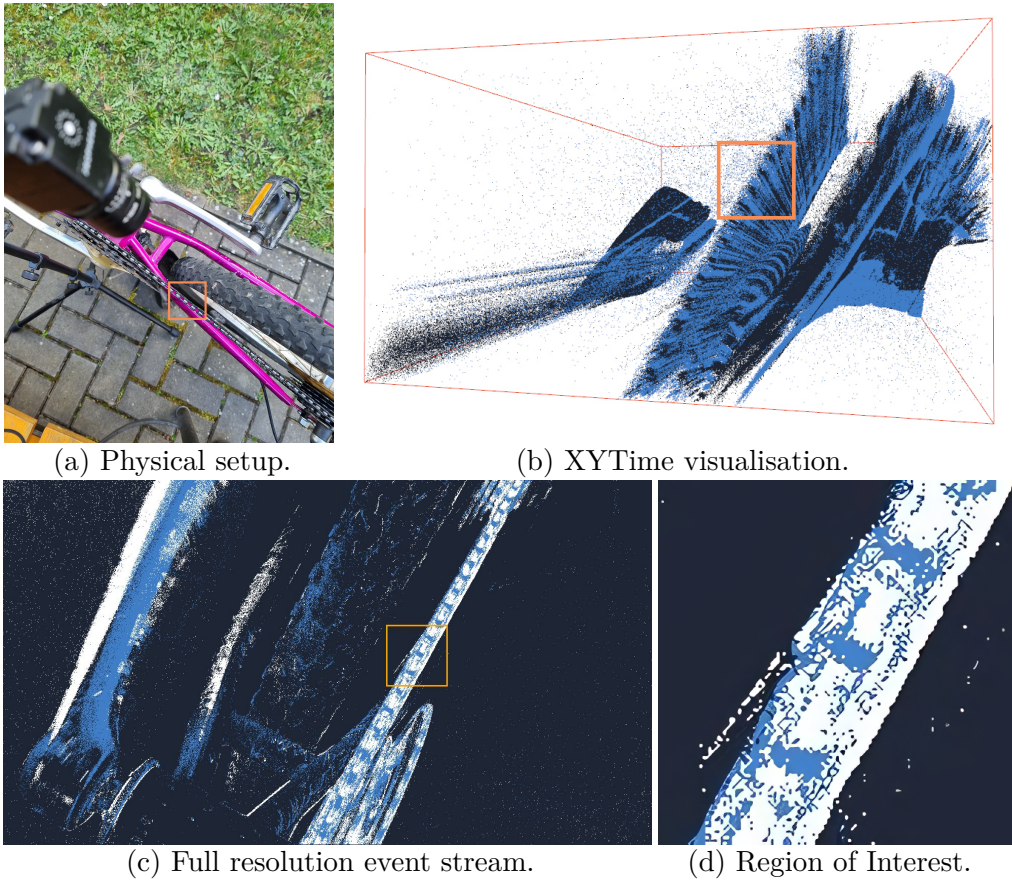(c) Full resolution event stream. (d) Region of Interest.

Figure 5.13: Bike chain from a top view (see Sec. 5.8.2): (a) physical setup for shooting the experiment; (b) events from a 250-millisecond window visualised in spatio-temporal space; (c) aggregated events captured by the entire sensor (1280×720px) within a time window equal to the period of the observed phenomenon with highlighted Region of Interest (180×180px) as shown in (d). Positive events are represented by white, and negative events are represented by bright blue.

Chapter **6**

Conclusion

In this thesis, we proposed a novel method of contactless measurement of periodic phenomena. The method performs a 3D correlation to detect periodic similarities within the event data stream of an event camera. Only a single assumption is made; the observed object periodically produces a similar set of events by returning to a known state or position.

We evaluated the performance of our method in twelve experiments involving objects that rotate, flash, vibrate, and move around, achieving a mean relative error of 0.172% across a wide range of frequencies (3.2 to 2000 Hz). We showed that our proposed method is suitable for measurement:

(i) in scenarios where objects lack prominent features, such as the rotating velcro disc experiment (see sections 5.5.2 and 5.5.3),

(ii) through transparent materials like glass or plastic as demonstrated in experiment 5.5.3,

(iii) of centrosymmetric objects (as shown in experiment 5.5.5), which can be challenging for some methods and leads them to detect a multiple of the ground truth frequency instead,

(iv) with non-frontal camera placement, offering flexibility in experimental setups, as demonstrated in experiments 5.5.3, 5.7.1, 5.5.5 and 5.8.2,

(v) regardless of the sensor or camera model used, including older models with lower resolutions (see experiments 5.5.4, 5.5.5 and 5.7.2),

(vi) from event stream with a significant amount of noise and slight camera instability as seen in experiment 5.5.6,

(vii) of frequences as high as 2000 Hz in experiment 5.6.1 (equivalent to 120000 RPM).

Appendix A

Used AI tools

Following *Methodical guideline No. 5/2023* [1], the following tools were used during the development of this thesis:

- Writefull[2] for spelling checking.

- Grammarly[3] for spelling and grammar checking.

- Amazon CodeWhisperer[4] for improved single-line code completion.

- DeepL[5] for occasional translations of English phrases into the Czech language and, in case of more complex English sentences, for creating draft phrasing proposals.
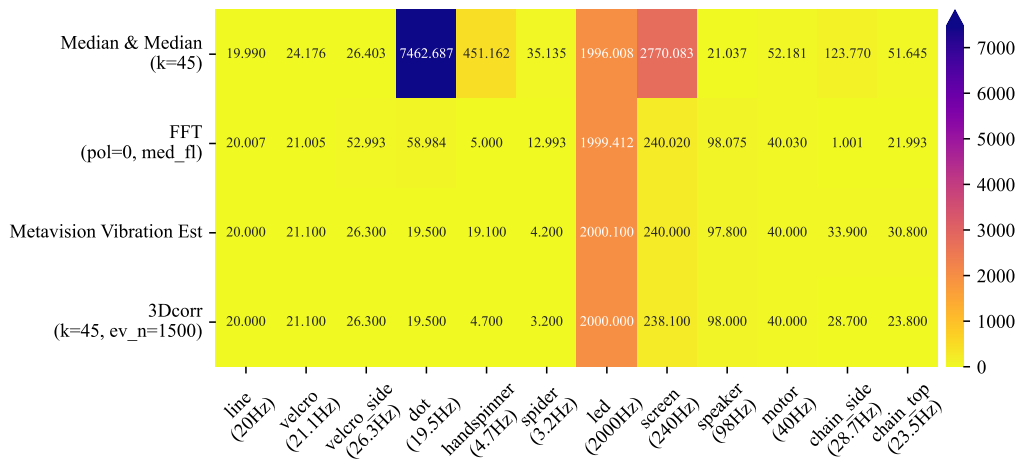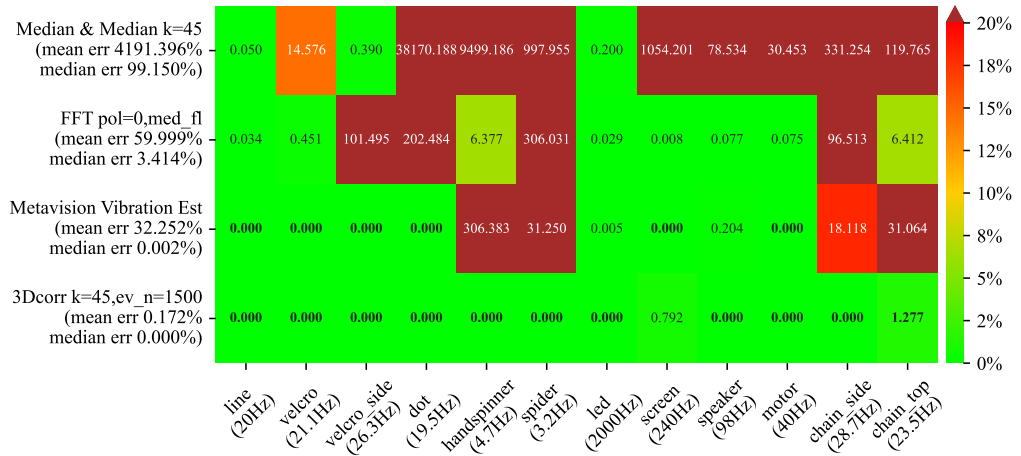
---

[1] https://www.cvut.cz/en/legislation-and-study-regulations#artificial-intelligence
[2] https://www.writefull.com
[3] https://www.grammarly.com
[4] https://aws.amazon.com/codewhisperer
[5] https://www.deepl.com/translator

Appendix B

Supplementary Materials

Method	line (20Hz)	velcro (21.1Hz)	velcro_side (26.3Hz)	dot (19.5Hz)	handspinner (4.7Hz)	spider (3.2Hz)	led (2000Hz)	screen (240Hz)	speaker (98Hz)	motor (40Hz)	chain_side (28.7Hz)	chain_top (23.5Hz)
Median & Median (k=45)	19.990	24.176	26.403	7462.687	451.162	35.135	1996.008	2770.083	21.037	52.181	123.770	51.645
FFT (pol=0, med_fl)	20.007	21.005	52.993	58.984	5.000	12.993	1999.412	240.020	98.075	40.030	1.001	21.993
Metavision Vibration Est	20.000	21.100	26.300	19.500	19.100	4.200	2000.100	240.000	97.800	40.000	33.900	30.800
3Dcorr (k=45, ev_n=1500)	20.000	21.100	26.300	19.500	4.700	3.200	2000.000	238.100	98.000	40.000	28.700	23.800

(a) Measured frequencies (Hz).

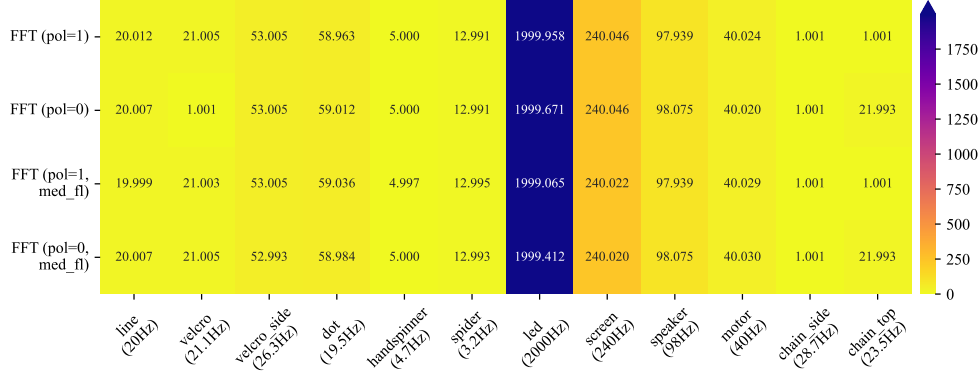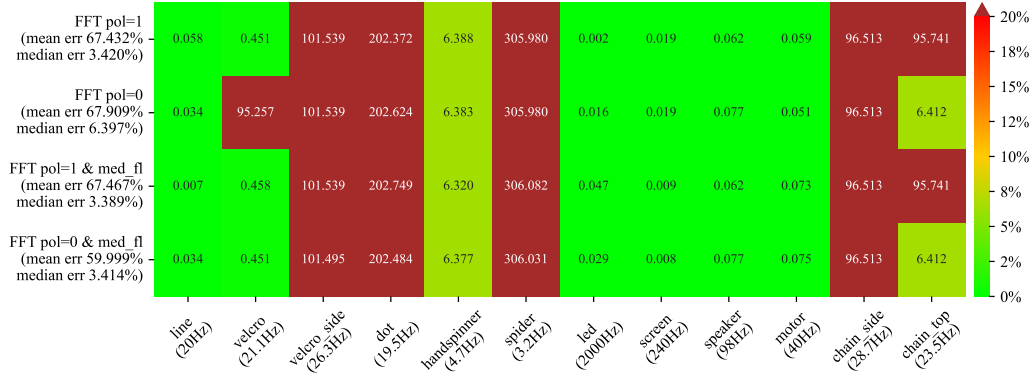| Method	line (20Hz)	velcro (21.1Hz)	velcro_side (26.3Hz)	dot (19.5Hz)	handspinner (4.7Hz)	spider (3.2Hz)	led (2000Hz)	screen (240Hz)	speaker (98Hz)	motor (40Hz)	chain_side (28.7Hz)	chain_top (23.5Hz)
Median & Median k=45 (mean err 4191.396% median err 99.150%)	0.050	14.576	0.390	38170.188	9499.186	997.955	0.200	1054.201	78.534	30.453	331.254	119.765
FFT pol=0,med_fl (mean err 59.999% median err 3.414%)	0.034	0.451	101.495	202.484	6.377	306.031	0.029	0.008	0.077	0.075	96.513	6.412
Metavision Vibration Est (mean err 32.252% median err 0.002%)	0.000	0.000	0.000	0.000	306.383	31.250	0.005	0.000	0.204	0.000	18.118	31.064
3Dcorr k=45,ev_n=1500 (mean err 0.172% median err 0.000%)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.792	0.000	0.000	0.000	1.277

(b) Relative errors (%).

Table B.1: Measured frequencies in hertz (a) and relative errors (b) achieved by the four methods (described in chapter 4) and twelve experiments (described in chapter 5). Optimal parameters for each method were found (see Sec. 5.2–5.4) and fixed across all experiments.

█ B.1 Parameter selection related materials

	line (20Hz)	velcro (21.1Hz)	velcro_side (26.3Hz)	dot (19.5Hz)	handspinner (4.7Hz)	spider (3.2Hz)	led (2000Hz)	screen (240Hz)	speaker (98Hz)	motor (40Hz)	chain_side (28.7Hz)	chain_top (23.5Hz)
FFT (pol=1)	20.012	21.005	53.005	58.963	5.000	12.991	1999.958	240.046	97.939	40.024	1.001	1.001
FFT (pol=0)	20.007	1.001	53.005	59.012	5.000	12.991	1999.671	240.046	98.075	40.020	1.001	21.993
FFT (pol=1, med_fl)	19.999	21.003	53.005	59.036	4.997	12.995	1999.065	240.022	97.939	40.029	1.001	1.001
FFT (pol=0, med_fl)	20.007	21.005	52.993	58.984	5.000	12.993	1999.412	240.020	98.075	40.030	1.001	21.993

(a) Measured frequencies (Hz).

	line (20Hz)	velcro (21.1Hz)	velcro_side (26.3Hz)	dot (19.5Hz)	handspinner (4.7Hz)	spider (3.2Hz)	led (2000Hz)	screen (240Hz)	speaker (98Hz)	motor (40Hz)	chain_side (28.7Hz)	chain_top (23.5Hz)
FFT pol=1 (mean err 67.432% median err 3.420%)	0.058	0.451	101.539	202.372	6.388	305.980	0.002	0.019	0.062	0.059	96.513	95.741
FFT pol=0 (mean err 67.909% median err 6.397%)	0.034	95.257	101.539	202.624	6.383	305.980	0.016	0.019	0.077	0.051	96.513	6.412
FFT pol=1 & med_fl (mean err 67.467% median err 3.389%)	0.007	0.458	101.539	202.749	6.320	306.082	0.047	0.009	0.062	0.073	96.513	95.741
FFT pol=0 & med_fl (mean err 59.999% median err 3.414%)	0.034	0.451	101.495	202.484	6.377	306.031	0.029	0.008	0.077	0.075	96.513	6.412

(b) Relative errors (%).

Table B.2: Measured frequencies in hertz (a) and relative errors (b) achieved by FFT analysis method (described in section 4.2) with different parameter configurations fixed across all experiments (described in chapter 5). The best results were achieved by analyzing negative events (`pol=0`) and applying median filtering (`med_pl`) – bottom row of (a) and (b). With this configuration, the method successfully measured the frequency in 6 out of 12 experiments with a relative error under 0.5%.
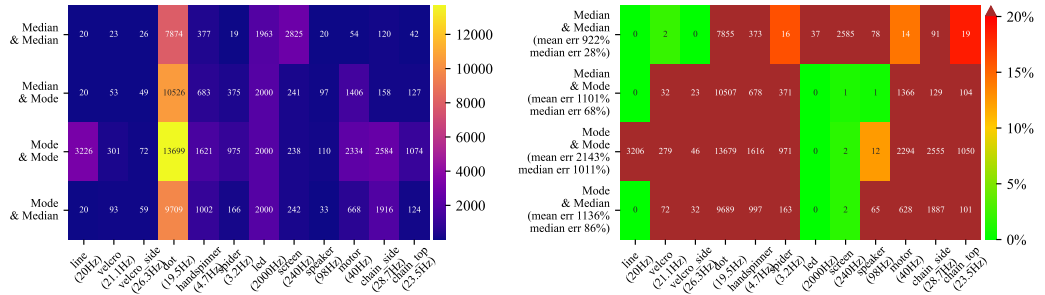
(a) Kernel size 30×30px.
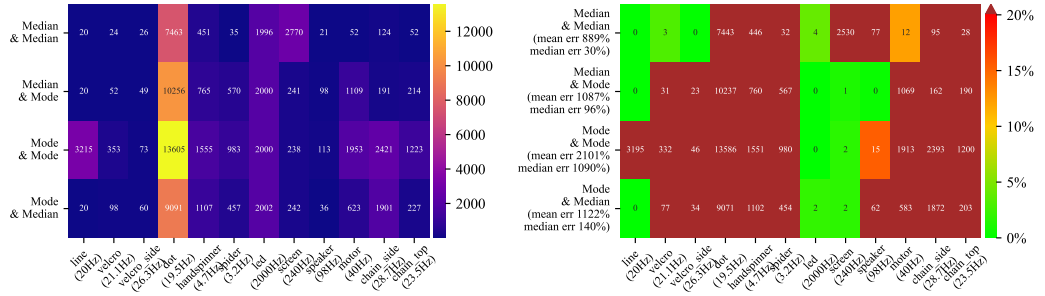
(b) Kernel size 45×45px.
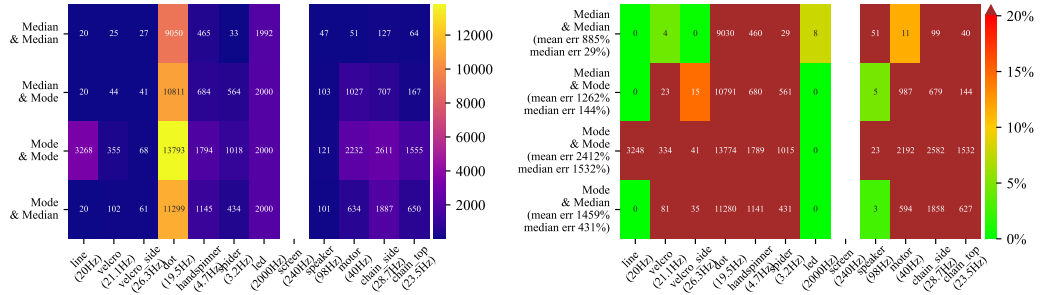
(c) Kernel size 60×60px.
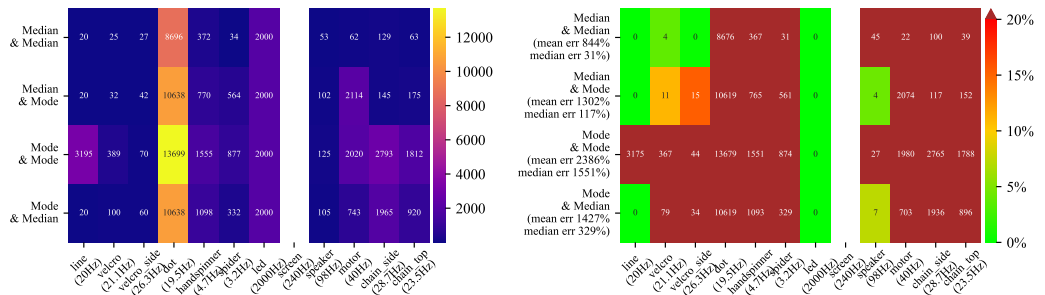
(d) Kernel size 75×75px.

Table B.3: Measured frequencies (left, Hz) and relative errors (right, %) of the baseline method (described in section 4.1) for four kernel sizes 30×30px (a), 45×45px (b), 60×60px (c), 75×75px (d). For each kernel size, all combinations of two temporal (median and mode) and two spatial analysis functions (median and mode) were examined and fixed in all experiments (described in chapter 5).

43
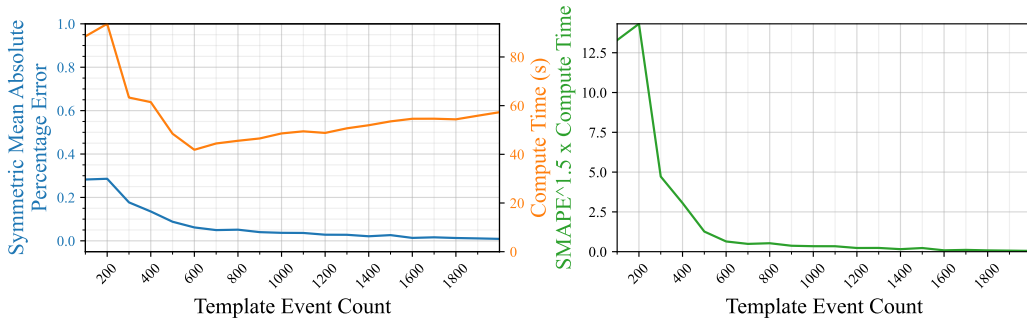
(a) Kernel size 30×30px.
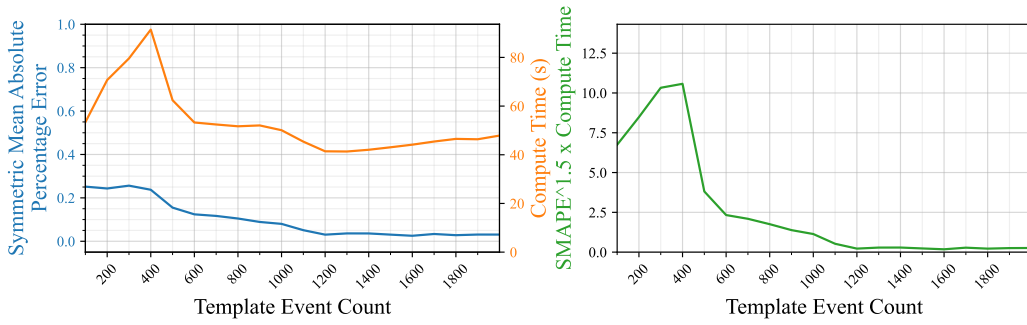
(b) Kernel size 45×45px.
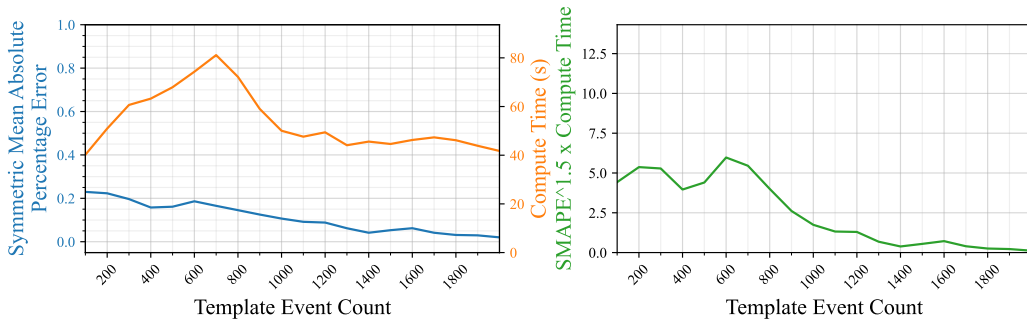
(c) Kernel size 60×60px.
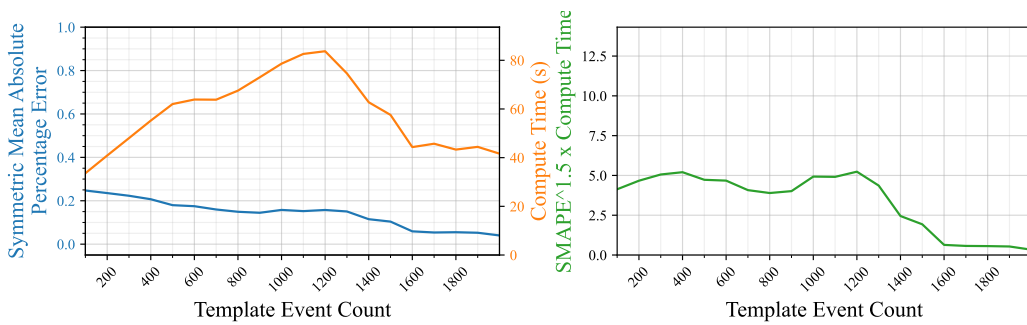
(d) Kernel size 75×75px.

Figure B.1: The side-view velcro disc experiment (see Sec. 5.5.3), SMAPE (5.1) values and computation time (left), and Performance-cost measure (right, see (5.2)) as a function of template event count for four kernel sizes 30×30px (a), 45×45px (b), 60×60px (c), 75×75px (d), see Sec. 5.4.
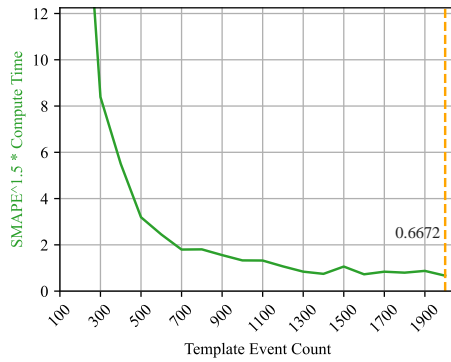
(a) Kernel size 30×30px.
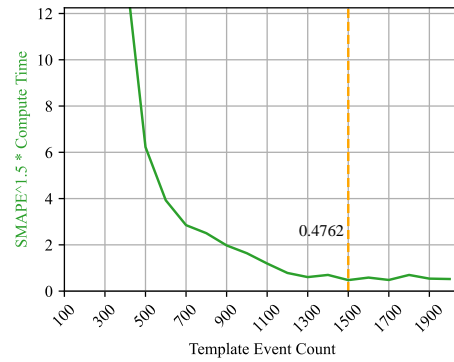
(b) Kernel size 45×45px.
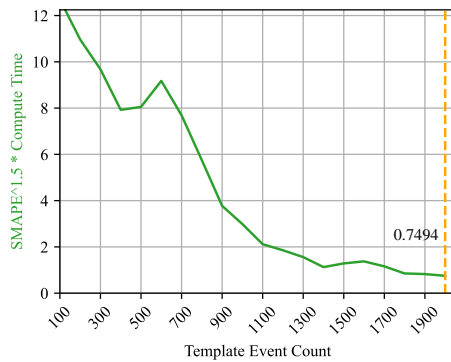
(c) Kernel size 60×60px.
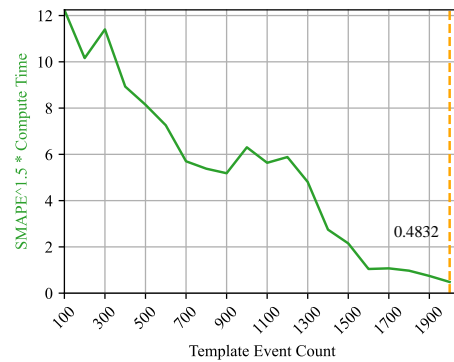
(d) Kernel size 75×75px.

Figure B.2: Performance-cost measure (5.2) of the 3D correlation method as a function of the template event count for four kernel sizes (30×30px (a), 45×45px (b), 60×60px (c), 75×75px (d)). The vertical orange lines indicate the best template event count for a specific kernel size. The best overall configuration is kernel size 45×45px with template count 1500, described in Sec. 5.4.

B.2   Individual experiments related materials

B.2.1   Felt disc with a high-contrast line

(a) Frequency relative error map.

(b) FFT amplitude map.

(c) Dominant frequencies.
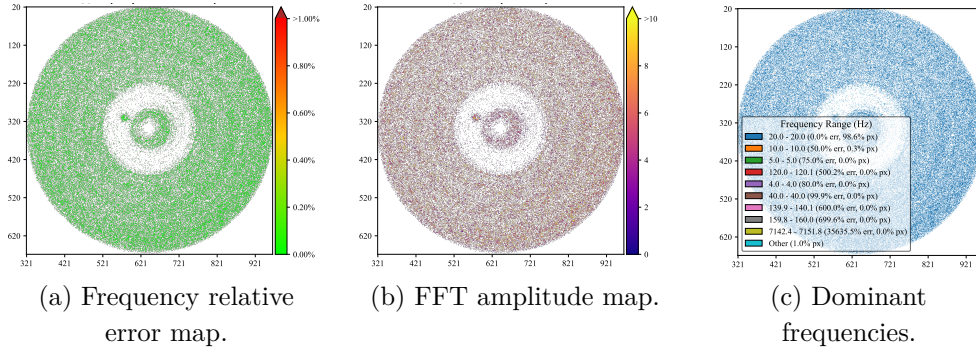
Figure B.3: FFT analysis method (see Sec. 4.2), experiment "Felt disc with a high-contrast line" (see Sec. 5.5.1): (a) per-pixel relative error in frequency estimation, (b) FFT amplitude map (higher values indicate a stronger presence of the specific frequency within the original signal), (c) pixel-wise classification based on the most frequent estimated frequencies.
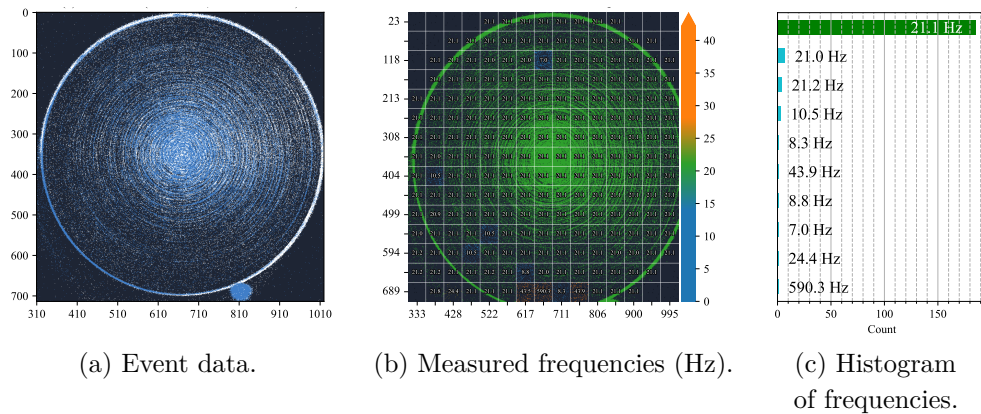

(a) Event data.

(b) Measured frequencies (Hz).

(c) Histogram of frequencies.
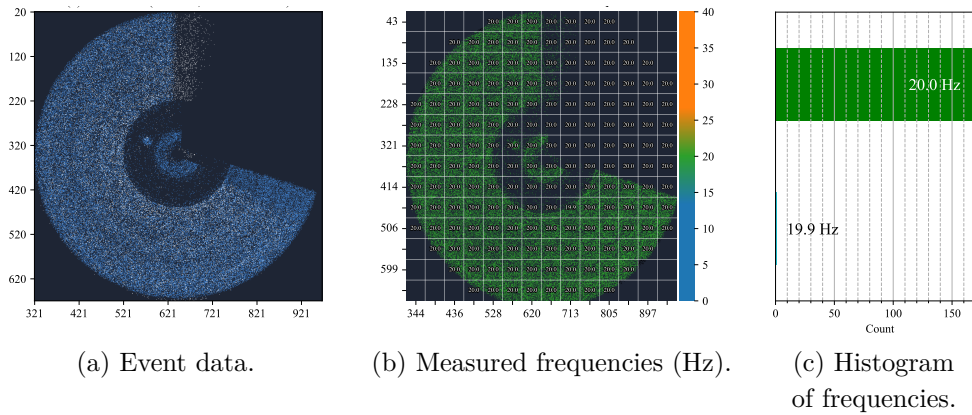
Figure B.4: 3D correlation method (see Sec. 4.3), experiment "Felt disc with a high-contrast line" experiment (see Sec. 5.5.1): (a) captured event data (white: positive event, blue: negative event), (b) frequency measurements across all 45×45px kernels (blue: underestimation, red: overestimation, grey: accurate/near-accurate estimation, black: insufficient events for analysis), (c) sorted histogram of measured frequencies.

B.2.2 Fronto-parallel velcro disc

(a) Frequency relative error map.

(b) FFT amplitude map.

(c) Dominant frequencies.
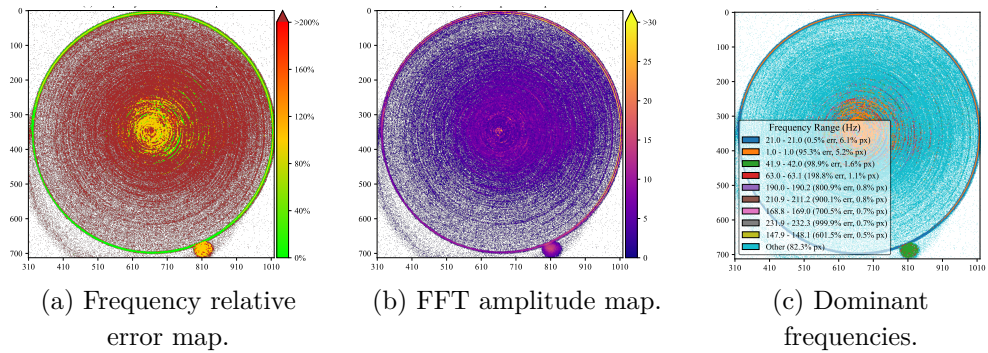
Figure B.5: FFT analysis method (see Sec. 4.2), experiment "Fronto-parallel velcro disc" (see Sec. 5.5.2): (a) per-pixel relative error in frequency estimation, (b) FFT amplitude map (higher values indicate a stronger presence of the specific frequency within the original signal), (c) pixel-wise classification based on the most frequent estimated frequencies.

(a) Event data.

(b) Measured frequencies (Hz).

(c) Histogram of frequencies.

Figure B.6: 3D correlation method (see Sec. 4.3), experiment "Fronto-parallel velcro disc" (see Sec. 5.5.2): (a) captured event data (white: positive event, blue: negative event), (b) frequency measurements across all 45×45px kernels (blue: underestimation, red: overestimation, grey: accurate/near-accurate estimation, black: insufficient events for analysis), (c) sorted histogram of measured frequencies.

47

■ B.2.3 Velcro disc with a non-frontal camera behind a glass sheet

(a) Frequency relative error map.

(b) FFT amplitude map.
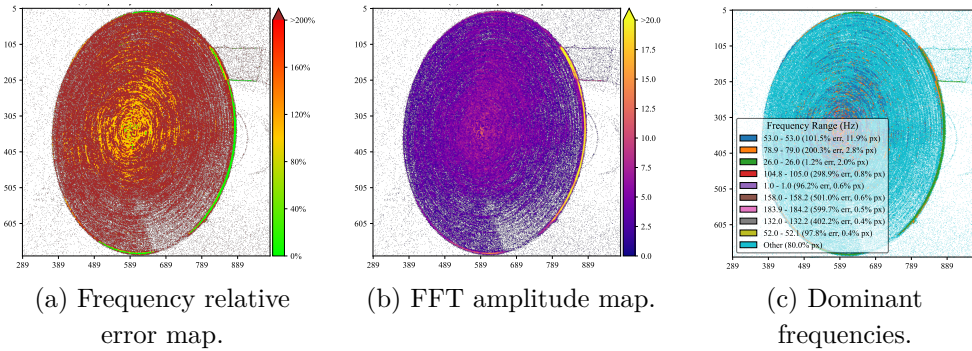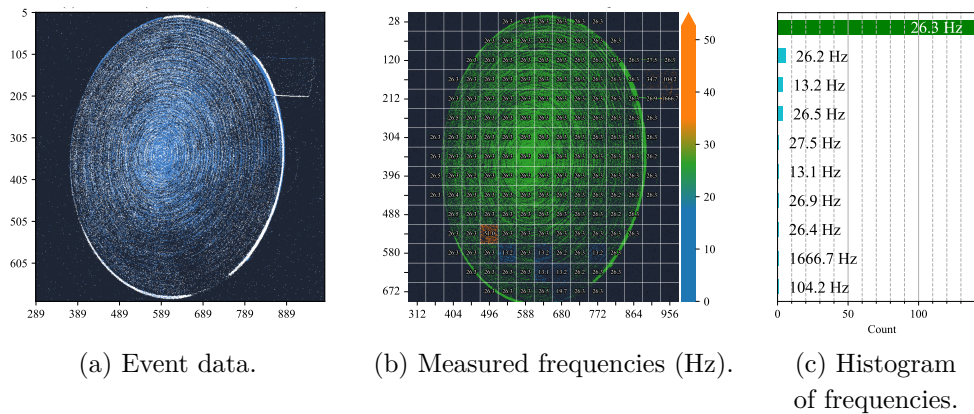
(c) Dominant frequencies.

Figure B.7: FFT analysis method (see Sec. 4.2), experiment "Velcro disc with a non-frontal camera behind a glass sheet" (see Sec. 5.5.3): (a) per-pixel relative error in frequency estimation, (b) FFT amplitude map (higher values indicate a stronger presence of the specific frequency within the original signal), (c) pixel-wise classification based on the most frequent estimated frequencies.

(a) Event data.

(b) Measured frequencies (Hz).

(c) Histogram of frequencies.

Figure B.8: 3D correlation method (see Sec. 4.3), experiment "Velcro disc with a non-frontal camera behind a glass sheet" (see Sec. 5.5.3): (a) captured event data (white: positive event, blue: negative event), (b) frequency measurements across all 45×45px kernels (blue: underestimation, red: overestimation, grey: accurate/near-accurate estimation, black: insufficient events for analysis), (c) sorted histogram of measured frequencies.

B.2.4 High-contrast dot

(a) Frequency relative error map.

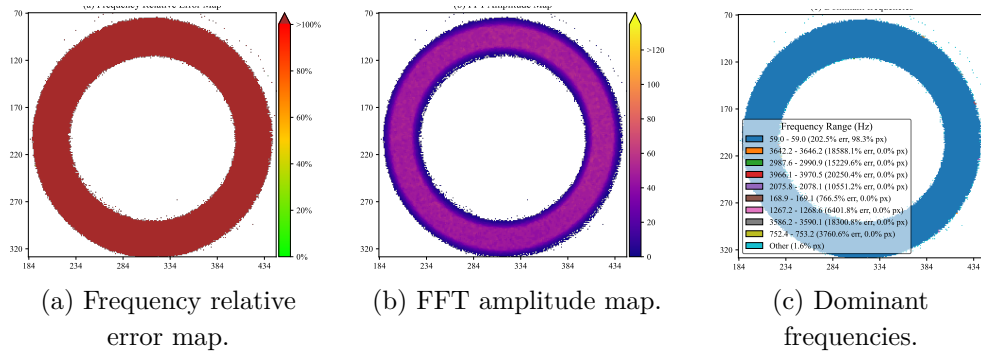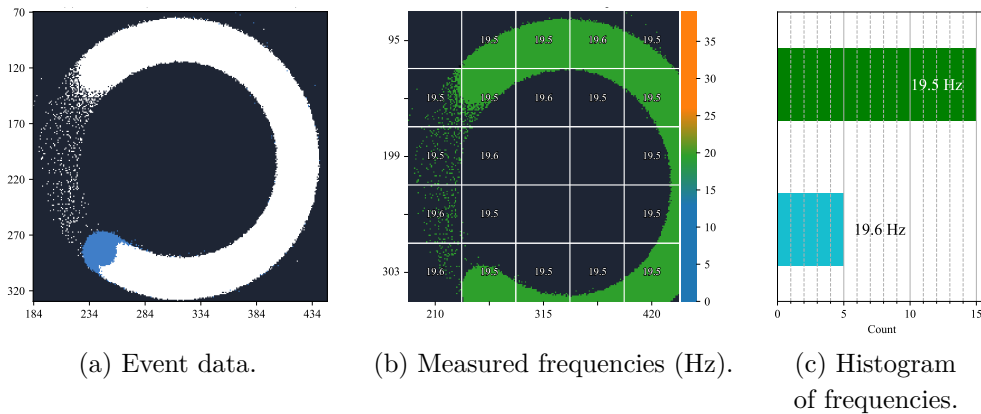(b) FFT amplitude map.

(c) Dominant frequencies.

Figure B.9: FFT analysis method (see Sec. 4.2), experiment "High-contrast dot" (see Sec. 5.5.4): (a) per-pixel relative error in frequency estimation, (b) FFT amplitude map (higher values indicate a stronger presence of the specific frequency within the original signal), (c) pixel-wise classification based on the most frequent estimated frequencies.

(a) Event data.

(b) Measured frequencies (Hz).

(c) Histogram of frequencies.

Figure B.10: 3D correlation method (see Sec. 4.3), experiment "High-contrast dot" (see Sec. 5.5.4): (a) captured event data (white: positive event, blue: negative event), (b) frequency measurements across all 45×45px kernels (blue: underestimation, red: overestimation, grey: accurate/near-accurate estimation, black: insufficient events for analysis), (c) sorted histogram of measured frequencies.

49

B.2.5 Hand fidget spinner

(a) Frequency relative error map.

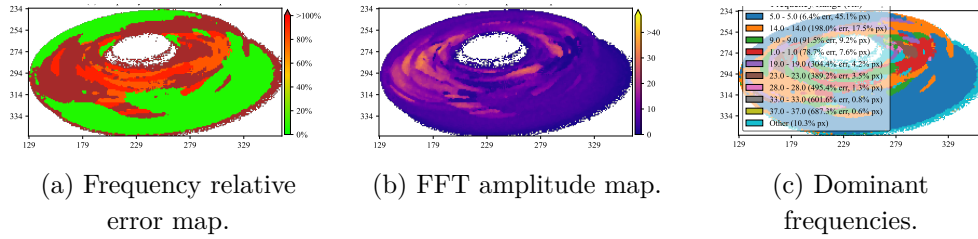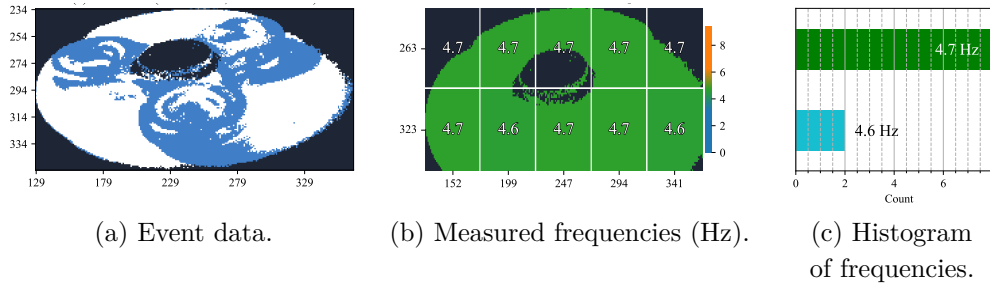(b) FFT amplitude map.

(c) Dominant frequencies.

Figure B.11: FFT analysis method (see Sec. 4.2), experiment "Hand fidget spinner" (see Sec. 5.5.5): (a) per-pixel relative error in frequency estimation, (b) FFT amplitude map (higher values indicate a stronger presence of the specific frequency within the original signal), (c) pixel-wise classification based on the most frequent estimated frequencies.

(a) Event data.

(b) Measured frequencies (Hz).

(c) Histogram of frequencies.

Figure B.12: 3D correlation method (see Sec. 4.3), experiment "Hand fidget spinner" (see Sec. 5.5.5): (a) captured event data (white: positive event, blue: negative event), (b) frequency measurements across all 45×45px kernels (blue: underestimation, red: overestimation, grey: accurate/near-accurate estimation, black: insufficient events for analysis), (c) sorted histogram of measured frequencies.

■ B.2.6   Whirling Pholcus phalangioides

(a) Frequency relative error map.

(b) FFT amplitude map.
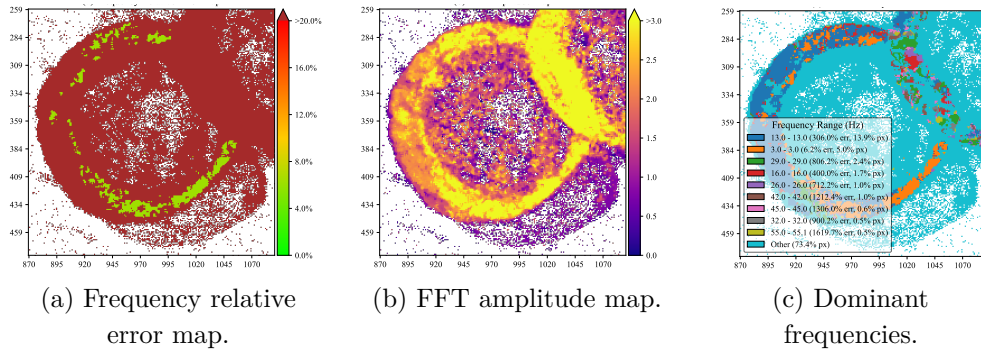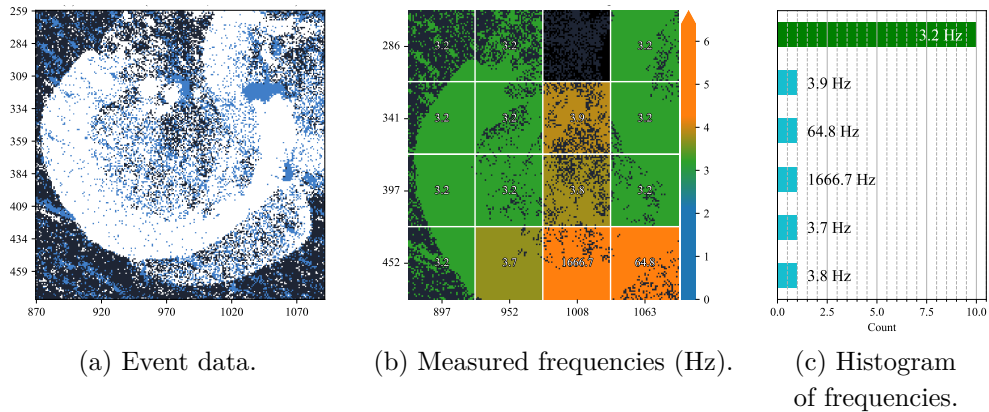
(c) Dominant frequencies.

Figure B.13: FFT analysis method (see Sec. 4.2), experiment "Whirling Pholcus phalangioides" (see Sec. 5.5.6): (a) per-pixel relative error in frequency estimation, (b) FFT amplitude map (higher values indicate a stronger presence of the specific frequency within the original signal), (c) pixel-wise classification based on the most frequent estimated frequencies.

(a) Event data.

(b) Measured frequencies (Hz).

(c) Histogram of frequencies.

Figure B.14: 3D correlation method (see Sec. 4.3), experiment "Whirling Pholcus phalangioides" (see Sec.5.5.6): (a) captured event data (white: positive event, blue: negative event), (b) frequency measurements across all 45×45px kernels (blue: underestimation, red: overestimation, grey: accurate/near-accurate estimation, black: insufficient events for analysis), (c) sorted histogram of measured frequencies.

■ **B.2.7** **Flashing LED**

(a) Frequency relative error map.

(b) FFT amplitude map.
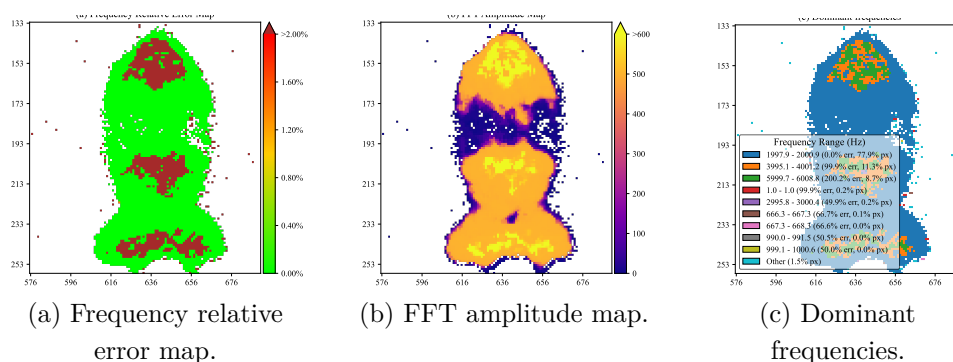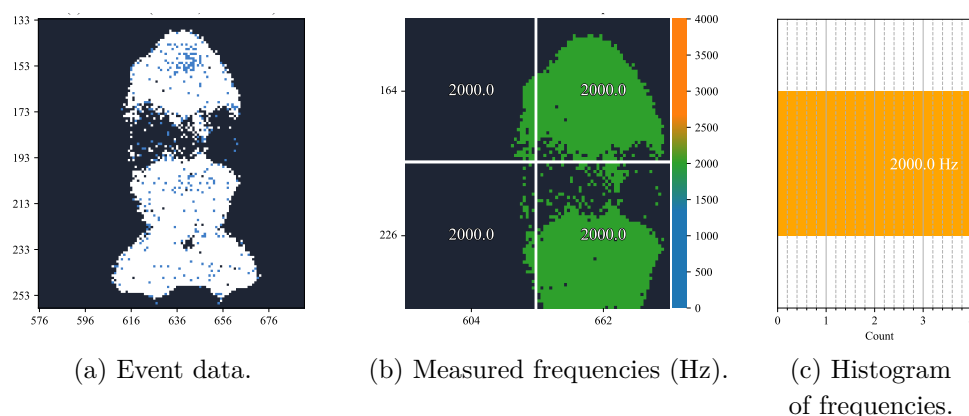
(c) Dominant frequencies.

Figure B.15: FFT analysis method (see Sec. 4.2), experiment "Flashing LED" (see Sec. 5.6.1): (a) per-pixel relative error in frequency estimation, (b) FFT amplitude map (higher values indicate a stronger presence of the specific frequency within the original signal), (c) pixel-wise classification based on the most frequent estimated frequencies.

(a) Event data.

(b) Measured frequencies (Hz).

(c) Histogram of frequencies.

Figure B.16: 3D correlation method (see Sec. 4.3), experiment "Flashing LED" (see Sec. 5.6.1): (a) captured event data (white: positive event, blue: negative event), (b) frequency measurements across all 45×45px kernels (blue: underestimation, red: overestimation, grey: accurate/near-accurate estimation, black: insufficient events for analysis), (c) sorted histogram of measured frequencies.

■ **B.2.8   Refreshing mobile phone screen**

(a) Frequency relative error map.
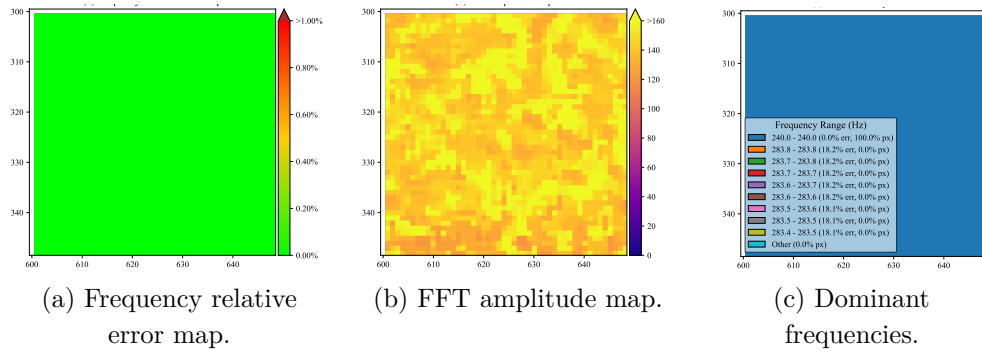
(b) FFT amplitude map.

(c) Dominant frequencies.

Figure B.17: FFT analysis method (see Sec. 4.2), experiment "Refreshing mobile phone screen" (see Sec. 5.6.2): (a) per-pixel relative error in frequency estimation, (b) FFT amplitude map (higher values indicate a stronger presence of the specific frequency within the original signal), (c) pixel-wise classification based on the most frequent estimated frequencies.
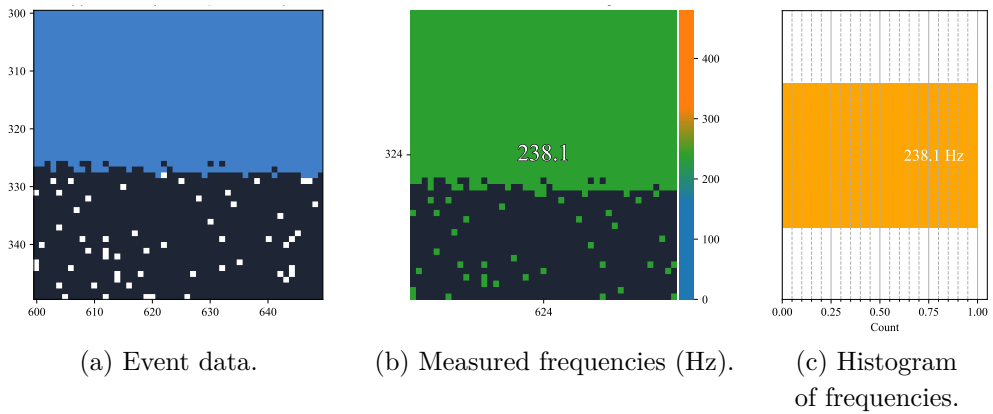

(a) Event data.

(b) Measured frequencies (Hz).

(c) Histogram of frequencies.

Figure B.18: 3D correlation method (see Sec. 4.3), experiment "Refreshing mobile phone screen" (see Sec. 5.6.2): (a) captured event data (white: positive event, blue: negative event), (b) frequency measurements across all 45×45px kernels (blue: underestimation, red: overestimation, grey: accurate/near-accurate estimation, black: insufficient events for analysis), (c) sorted histogram of measured frequencies.

■ **B.2.9** **Speaker diaphragm**

(a) Frequency relative error map.

(b) FFT amplitude map.
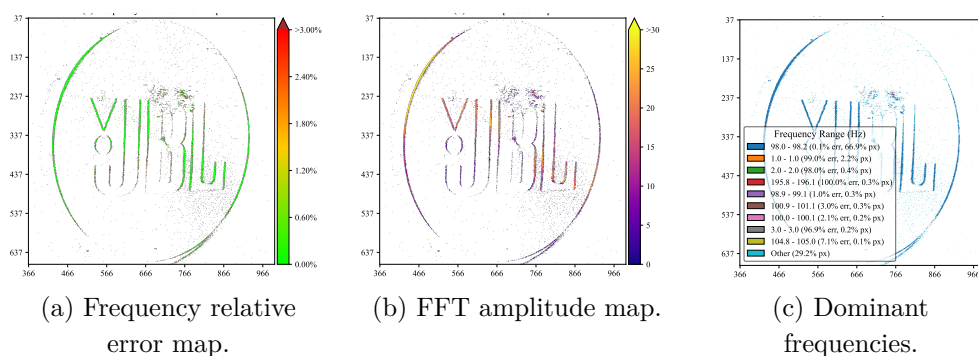
(c) Dominant frequencies.

Figure B.19: FFT analysis method (see Sec. 4.2), experiment "Speaker diaphragm" (see Sec. 5.7.1): (a) per-pixel relative error in frequency estimation, (b) FFT amplitude map (higher values indicate a stronger presence of the specific frequency within the original signal), (c) pixel-wise classification based on the most frequent estimated frequencies.
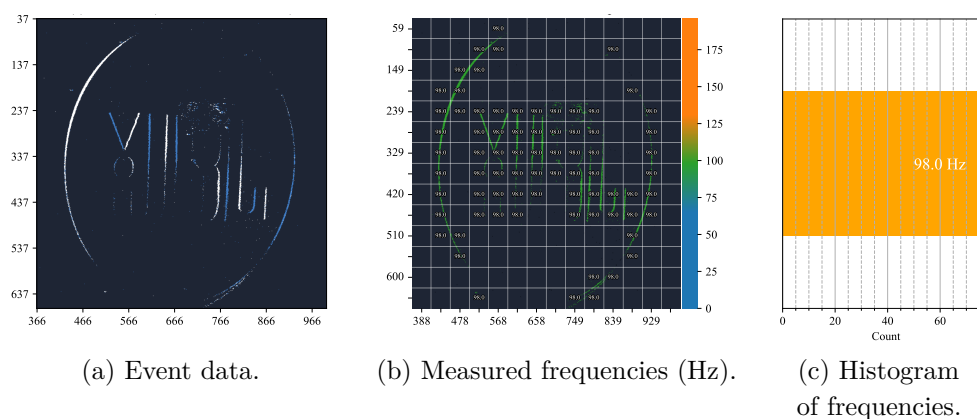

(a) Event data.

(b) Measured frequencies (Hz).

(c) Histogram of frequencies.

Figure B.20: 3D correlation method (see Sec. 4.3), experiment "Speaker diaphragm" (see Sec. 5.7.1): (a) captured event data (white: positive event, blue: negative event), (b) frequency measurements across all 45×45px kernels (blue: underestimation, red: overestimation, grey: accurate/near-accurate estimation, black: insufficient events for analysis), (c) sorted histogram of measured frequencies.

▪ **B.2.10** **Vibrating motor**

(a) Frequency relative error map.

(b) FFT amplitude map.
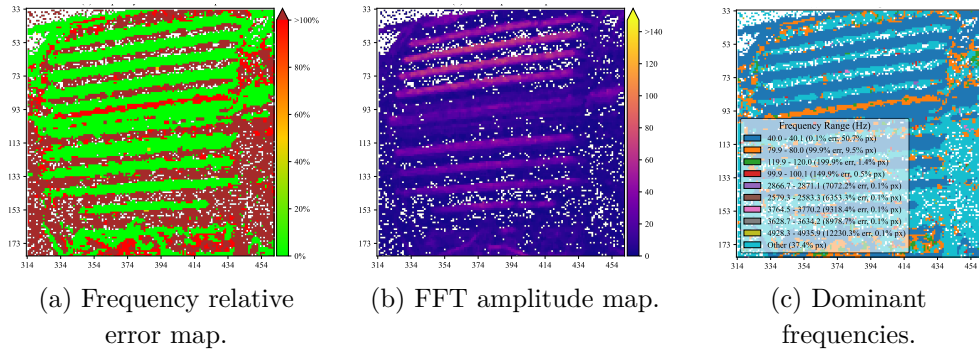
(c) Dominant frequencies.

Figure B.21: FFT analysis method (see Sec. 4.2), experiment "Vibrating motor" (see Sec. 5.7.2): (a) per-pixel relative error in frequency estimation, (b) FFT amplitude map (higher values indicate a stronger presence of the specific frequency within the original signal), (c) pixel-wise classification based on the most frequent estimated frequencies.
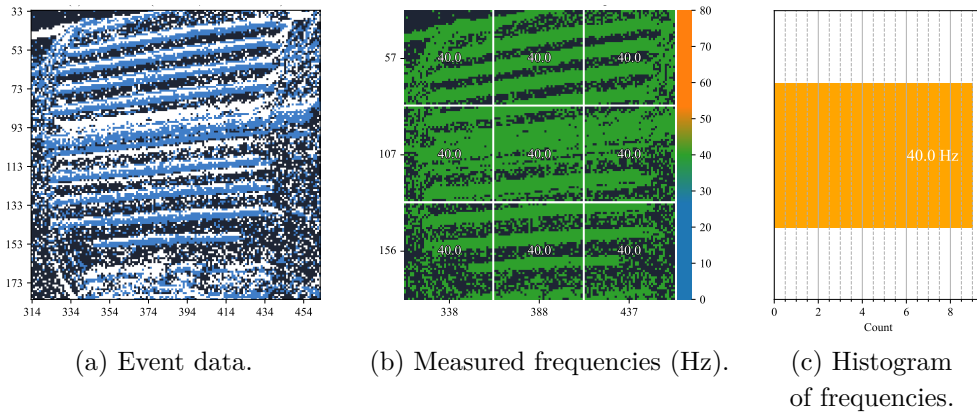

(a) Event data.

(b) Measured frequencies (Hz).

(c) Histogram of frequencies.

Figure B.22: 3D correlation method (see Sec. 4.3), experiment "Vibrating motor" (see Sec. 5.7.2): (a) captured event data (white: positive event, blue: negative event), (b) frequency measurements across all 45×45px kernels (blue: underestimation, red: overestimation, grey: accurate/near-accurate estimation, black: insufficient events for analysis), (c) sorted histogram of measured frequencies.

■ B.2.11 Bike chain from side view

(a) Frequency relative
error map.

(b) FFT amplitude map.
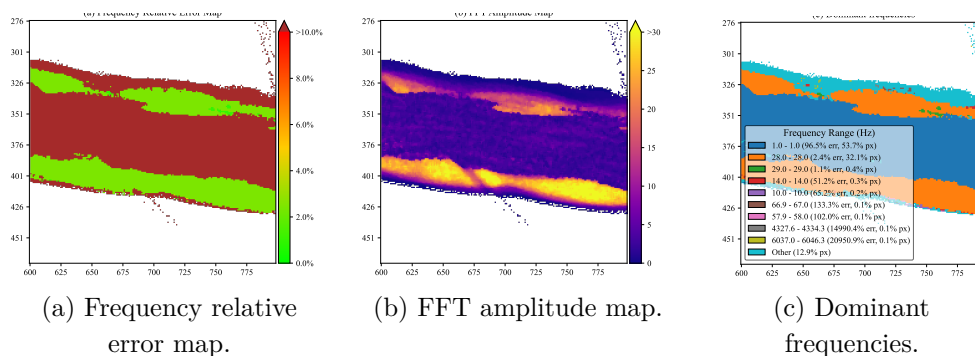
(c) Dominant
frequencies.

Figure B.23: FFT analysis method (see Sec. 4.2), experiment "Bike chain from side view" (see Sec. 5.8.1): (a) per-pixel relative error in frequency estimation, (b) FFT amplitude map (higher values indicate a stronger presence of the specific frequency within the original signal), (c) pixel-wise classification based on the most frequent estimated frequencies.
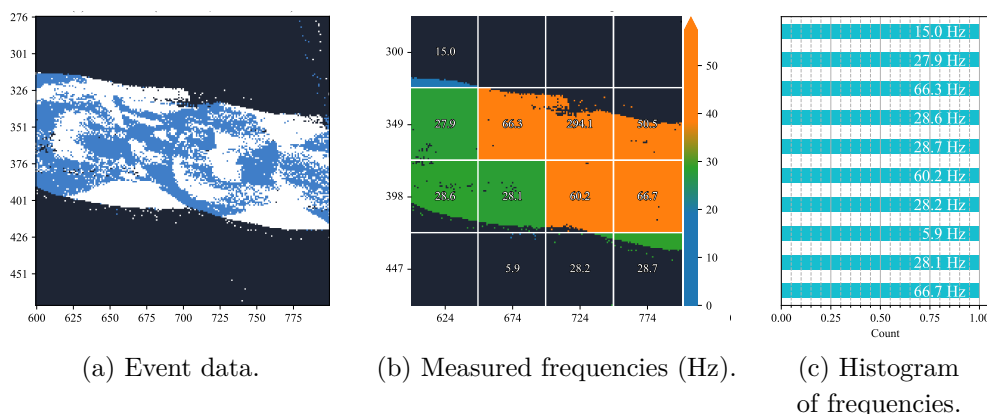

(a) Event data.

(b) Measured frequencies (Hz).

(c) Histogram
of frequencies.

Figure B.24: 3D correlation method (see Sec. 4.3), experiment "Bike chain from side view" (see Sec. 5.8.1): (a) captured event data (white: positive event, blue: negative event), (b) frequency measurements across all 45×45px kernels (blue: underestimation, red: overestimation, grey: accurate/near-accurate estimation, black: insufficient events for analysis), (c) sorted histogram of measured frequencies.

56

■ B.2.12 Bike chain from top view

(a) Frequency relative error map.

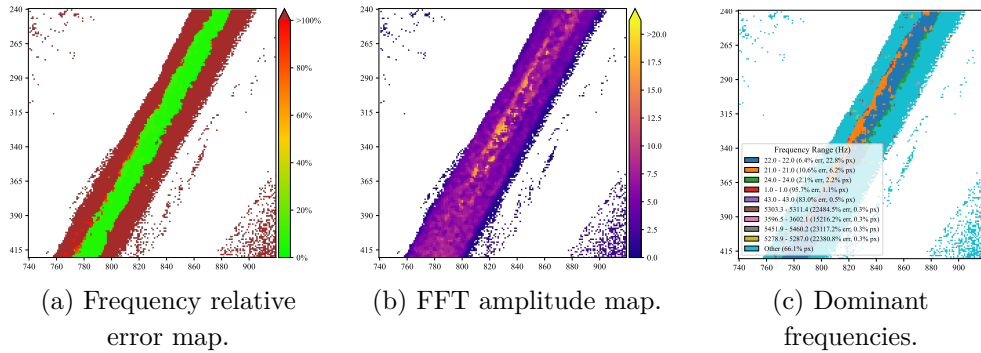(b) FFT amplitude map.

(c) Dominant frequencies.

Figure B.25: FFT analysis method (see Sec.4.2), experiment "Bike chain from top view" (see Sec.5.8.2): (a) per-pixel relative error in frequency estimation, (b) FFT amplitude map (higher values indicate a stronger presence of the specific frequency within the original signal), (c) pixel-wise classification based on the most frequent estimated frequencies.
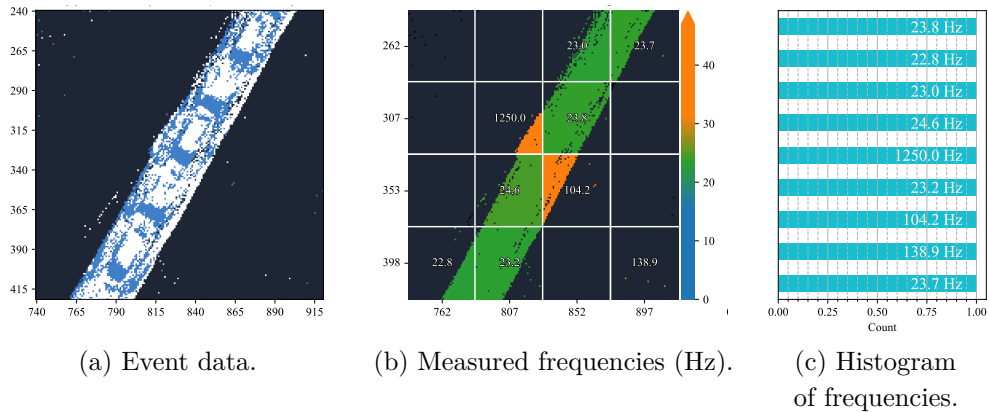

(a) Event data.

(b) Measured frequencies (Hz).

(c) Histogram of frequencies.

Figure B.26: 3D correlation method (see Sec. 4.3), experiment "Bike chain from top view" (see Sec. 5.8.2): (a) captured event data (white: positive event, blue: negative event), (b) frequency measurements across all 45 × 45px kernels (blue: underestimation, red: overestimation, grey: accurate/near-accurate estimation, black: insufficient events for analysis), (c) sorted histogram of measured frequencies.

Appendix C

Bibliography

[1] Electrical Safety Testing Laboratory, Institute of Testing & Certification India Pvt. Ltd, "Motor torque and speed testing," May 2023.

[2] A. Singh and Y. Kim, "Accurate measurement of drone's blade length and rotation rate using pattern analysis with W-band radar," *Electronics Letters*, vol. 54, pp. 523–525, Apr. 2018.

[3] P. R. Kamble, A. G. Keskar, and K. M. Bhurchandi, "Ball tracking in sports: a survey," *Artificial Intelligence Review*, vol. 52, pp. 1655–1705, 2019.

[4] Thunder Said Energy, "How is the power of a wind turbine calculated? - Thunder Said," Nov. 2022.

[5] Uni-Trend Technology (China) Co., Ltd., "UT370 Series Tachometers - UNI-T Meters | Test & Measurement Tools and Solutions," Aug. 2022.

[6] M. H. Mohd Ghazali and W. Rahiman, "Vibration Analysis for Machine Monitoring and Diagnosis: A Systematic Review," *Shock and Vibration*, vol. 2021, pp. 1–25, Sept. 2021.

[7] Prophesee S.A., "Prophesee Showcases Event Based Vision Partners at VISION 2022," Oct. 2022.

[8] E. Wallace, "Machine Vision Plays Crucial Role Modern Manufacturing," Feb. 2024.

[9] R. Tapia, J. P. Rodríguez-Gómez, J. A. Sanchez-Diaz, F. J. Gañán, I. G. Rodríguez, J. Luna-Santamaria, J. R. M. de Dios, and A. Ollero, "A comparison between frame-based and event-based cameras for flapping-wing robot perception," 2023.

[10] A. Jianyi and A. Dasgupta, "Frame and Event-Based Neural Network based Vision Sensors," preprint, Open Science Framework, Apr. 2023.

[11] J. P. Boettiger, "A Comparative Evaluation of the Detection and Tracking Capability Between Novel Event-Based and Conventional Frame-Based Sensors," *Theses and Dissertations*, Mar. 2020.

[12] T. Finateu, A. Niwa, D. Matolin, K. Tsuchimoto, A. Mascheroni, E. Reynaud, P. Mostafalu, F. Brady, L. Chotard, F. LeGoff, H. Takahashi, H. Wakabayashi, Y. Oike, and C. Posch, "5.10 A 1280×720 Back-Illuminated Stacked Temporal Contrast Event-Based Vision Sensor with 4.86µm Pixels, 1.066GEPS Readout, Programmable Event-Rate Controller and Compressive Data-Formatting Pipeline," in *2020 IEEE International Solid-State Circuits Conference - (ISSCC)*, pp. 112–114, 2020.

[13] S. Corporation, "Sony to Release Two Types of Stacked Event-Based Vision Sensors with the Industry's Smallest 4.86um Pixel Size for Detecting Subject Changes Only," Sept. 2021.

[14] Prophesee S.A., "Biases — Metavision SDK Docs 4.5.0 documentation," Dec. 2022.

[15] G. Chen, H. Cao, M. Aafaque, J. Chen, C. Ye, F. Röhrbein, J. Conradt, K. Chen, Z. Bing, X. Liu, G. Hinz, W. Stechele, and A. Knoll, "Neuromorphic Vision Based Multivehicle Detection and Tracking for Intelligent Transportation System," *Journal of Advanced Transportation*, vol. 2018, pp. 1–13, Dec. 2018.

[16] L. Li, H. Hu, Y. Qin, and K. Tang, "Digital approach to rotational speed measurement using an electrostatic sensor," *Sensors*, vol. 19, no. 11, 2019.

[17] H. Austerlitz, "Chapter 2 - analog signal transducers," in *Data Acquisition Techniques Using PCs (Second Edition)* (H. Austerlitz, ed.), pp. 6–28, San Diego: Academic Press, second edition ed., 2003.

[18] Metravi Instruments, "How do Digital Non-contact Tachometers work?."

[19] T. Wang, Y. Yan, L. Wang, and Y. Hu, "Rotational speed measurement through image similarity evaluation and spectral analysis," *IEEE Access*, vol. 6, pp. 46718–46730, 2018.

[20] Y. Wang, L. Wang, and Y. Yan, "Rotational speed measurement through digital imaging and image processing," in *2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pp. 1–6, 2017.

[21] F. Natili, F. Castellani, D. Astolfi, and M. Becchetti, "Video-Tachometer Methodology for Wind Turbine Rotor Speed Measurement," *Sensors*, vol. 20, no. 24, 2020.

[22] K. W. Hylton, P. Mitchell, B. Van Hoy, and T. P. Karnowski, "Experiments and analysis for measuring mechanical motion with event cameras," *Electronic Imaging*, vol. 33, pp. 333–1–333–8, Jan. 2021.

[23] G. O. d. A. Azevedo, B. J. T. Fernandes, L. H. d. S. Silva, A. Freire, R. P. de Araújo, and F. Cruz, "Event-Based Angular Speed Measurement and Movement Monitoring," *Sensors (Basel, Switzerland)*, vol. 22, p. 7963, Oct. 2022.

[24] G. Zhao, Y. Shen, N. Chen, P. Hu, L. Liu, and H. Wen, "High Speed Rotation Estimation with Dynamic Vision Sensors," Sept. 2022.

[25] C. Shi, N. Song, B. Wei, Y. Li, Y. Zhang, W. Li, and J. Jin, "Event-based vibration frequency measurement with laser-assisted illumination based on mixture gaussian distribution," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–13, 2023.

[26] J. Kolář, R. Špetlík, and J. Matas, "Ee3p: Event-based estimation of periodic phenomena properties," *arXiv preprint arXiv:2402.14958*, 2024.

[27] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex fourier series," *Mathematics of Computation*, vol. 19, pp. 297–301, 1965.

[28] Prophesee S.A., "Vibration Estimation using Python — Metavision SDK Docs 4.5.2 documentation."

[29] Prophesee S.A., "Recordings and Datasets — Metavision SDK Docs 4.5.2 documentation."

[30] Prophesee S.A., "Inference Pipeline of Event to Video — Metavision SDK Docs 4.5.2 documentation."

[31] R. R. Jackson, R. J. Brassington, and R. J. Rowe, "Anti-predator defences of Pholcus phalangioides (Araneae, Pholcidae), a web-building and web-invading spider," *Journal of Zoology*, vol. 220, no. 4, pp. 543–552, 1990.

[32] J. Fiala, "Raspberry Pi Pico oscilloscope with a web-based user interface," Master's thesis, České vysoké učení technické v Praze, Praha, June 2023.

[33] D. Schmidt, "Samsung Galaxy S21 Ultra Smartphone Review: Improved in many ways, but not all," Jan. 2021.

[34] LuxDeLux, "Frequency Sound Generator," Sept. 2023.