Seeing like a Cephalopod: Colour Vision with a Monochrome Event Camera -Supplementary Materials-

Sami Arja [1], Nimrod Kruger, Alexandre Marcireau, Nicholas Owen Ralph
Saeed Afshar and Gregory Cohen

In this document, we describe our synchronisation, light source calibration approach used in the paper "Seeing like a Cephalopod: Colour Vision with a Monochrome Event Camera" and show a simple technique to perform colour segmentation by focusing.

I. Synchronisation and Data Acquisition Hardware

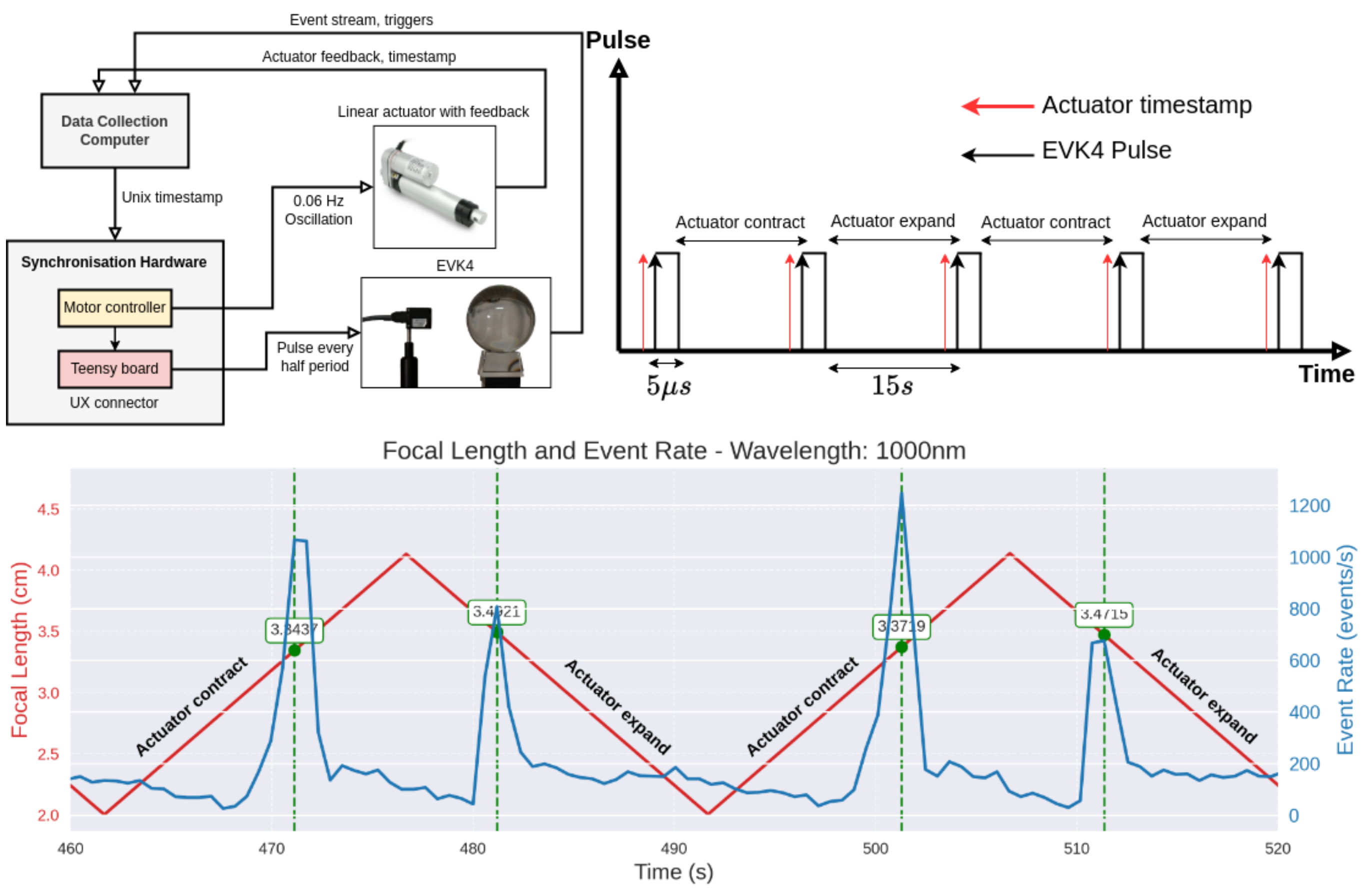

Fig. 1: The system architecture for data acquisition. **Top:** The event camera is synchronised with the linear actuator to obtain accurate focal distance values. **Bottom:** The measured event rate (blue) and focal distance (red) curves demonstrate the consistent peak in event rate over the same focal point.

Due to the event camera's high temporal resolution, accurate synchronisation is essential. In this study, synchronisation was achieved using a physical synchronisation signal between the host computer, linear actuator, and event camera (Figure 1).

The event camera's synchronisation cable (IX connector) connects to a Teensy board, which generates trigger pulses (rising and falling edges) each time the linear actuator completes a half oscillation ($\frac{f}{2}$). Since the Teensy lacks a Real-Time Clock (RTC), the host computer provides a Unix timestamp to the Teensy at the start of actuator movement. Both the timestamp and actuator feedback (in mm) are recorded.

[1] S. Arja, N. Kruger, A. Alexander, N. Ralph, S. Afshar, G. Cohen are with the International Centre for Neuromorphic Systems, Western Sydney University, Australia, s.elarja@westernsydney.edu.au

The actuator was programmed to oscillate between 1 cm and 3 cm at 0.06 Hz. At each maximum extension and contraction, a synchronization pulse is sent to the event camera, and the corresponding Unix timestamp and actuator position are logged. Event data acquisition was performed using Gen4 software[1], which automatically records trigger timestamps.

Due to initial misalignment between host computer and event camera timestamps, post-processing alignment was performed. We determined the consistent offset between trigger events and actuator timestamps and corrected all timestamps accordingly. The synchronization delay was approximately 10–50 ms, acceptable given the slow actuator frequency. The low frequency was chosen specifically because the event camera's response slows in dark environments, necessitating slower movements to generate sufficient event data.

Figure 1 (bottom panel) demonstrates event rate peaks during actuator contraction and expansion, clearly indicating that at these focal positions, the camera focuses near-infrared (1000 nm) wavelengths while other wavelengths remain defocused.

II. Light Source Calibration

Because the camera spectral sensitivity depends on the quantum efficiency, the wavelength intensity from the monochromator [2] needs to be adjusted to produce the same number of photons for the imaging sensor. The raw intensity (i.e. frame) is captured by the camera at a given wavelength λ that is proportional to the product of the incident power $P(\lambda)$ and the QE, $\mathrm{QE}(\lambda)$. Because $\mathrm{QE}(\lambda)$ of both the Sony IMX249 [3] and the EVK4 decay rapidly after 700 nm toward 1000 nm, directly measuring the same optical power at all wavelengths would result in a camera signal biased by the camera's spectral response (i.e. Quantum efficiency). To address this, we designate 1000 nm (where the camera is least sensitive) as our reference. That is, we measure the "full-iris" power at 1000 nm, $P_{\text{full}}(1000)$, and multiply by $\mathrm{QE}(1000)$ to determine the baseline camera signal for the dimmest region of the spectrum. By using this low-sensitivity endpoint as the reference, we avoid overexposure or saturation at shorter wavelengths. We then replicate this same "reference signal" at each other wavelength λ by adjusting the incident power such that

$$P_{\text{target}}(\lambda) \times \mathrm{QE}(\lambda) = P_{\text{full}}(1000) \times \mathrm{QE}(1000).$$

In practice, we also measure and subtract a small "dark power" offset, P_{dark}, to account for baseline noise in our system. Specifically, the net photoelectron-generating signal becomes

$$\Big[P_{\text{full}}(\lambda) \times \mathrm{QE}(\lambda)\Big] - P_{\text{dark}},$$

and we adjust $P_{\text{full}}(\lambda)$ accordingly to match that obtained at 1000 nm. By calibrating in this manner, the camera captures the same effective brightness at every wavelength, thereby ensuring meaningful, comparable data across the visible and near-infrared range. Further details, including the measured, adjusted, and final target power values, are provided in Table I.

TABLE I: Light source calibration values across visible and near-infrared.

λ (nm)	QE	P_{full} (nW)	P_{dark} (nW)	P_{adjusted} = (P_{full}/QE)-P_{dark} (nW)	P_{target} (nW)
400	0.5805	36.28	0.02554	21.04	5.43
450	0.7422	84.28	0.02554	62.53	4.24
500	0.7852	147.4	0.02554	115.72	4.01
550	0.7060	184.2	0.02554	130.03	4.46
600	0.6088	198.2	0.02554	120.65	5.17
650	0.5008	195.2	0.02554	97.74	6.29
700	0.3892	164.2	0.02554	63.89	8.09
750	0.2884	133.1	0.02554	38.37	10.92
800	0.1966	116.9	0.02554	22.97	16.02
850	0.1318	109.1	0.02554	14.36	23.89
900	0.0795	135.6	0.02554	10.77	39.58
950	0.0399	194.1	0.02554	7.72	78.93
1000	0.0146	217.1	0.02554	3.15	215.39

III. Longitudinal Chromatic Aberration of the Ball Lens

We investigated the type of chromatic aberration produced by the ball lens, focusing specifically on distinguishing between Longitudinal Chromatic Aberration (LCA, or Axial Chromatic Aberration) and Transverse Chromatic Aberration (Lateral Chromatic Aberration). Our experiment confirmed that the ball lens primarily produces Longitudinal Chromatic Aberration, where different wavelengths focus at distinct points along the optical axis, resulting in uniform aberration across the entire field of view regardless of the point source's position.

[1] https://github.com/neuromorphicsystems/gen4
[2] https://www.newport.com/f/cs130b-configured-monochromators
[3] https://en.ids-imaging.com/store/u3-3262se-rev-1-2.html

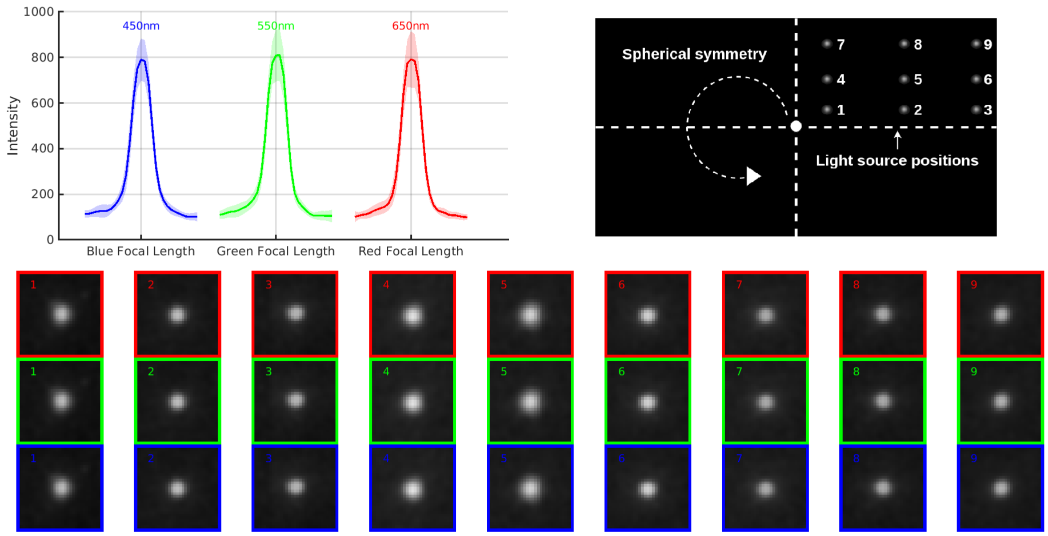

Fig. 2: Longitudinal chromatic aberration from the ball lens. Different points across the field of view produce similar chromatic aberrations effects, enabling consistent spectral discrimination at any position within the image.

To validate this, we moved a pinhole to nine distinct positions within the first quadrant of the imaging field of view. Due to the spherical symmetry of the ball lens, the aberration pattern observed in one quadrant is identical in the other quadrants. Figure 2 illustrates the presence of LCA. For each position, we separately recorded images at three wavelengths (450 nm, 550 nm, and 650 nm) while adjusting the focal plane accordingly for each wavelength. A total of nine images per position (three wavelengths, each with three focal planes) were captured. The aggregated cross-sectional profiles for each wavelength, shown in Figure 2 (top left), highlight subtle yet consistent differences between wavelengths and demonstrate the lens's uniform aberration characteristics.

IV. Spectral (Colour) Segmentation Algorithm

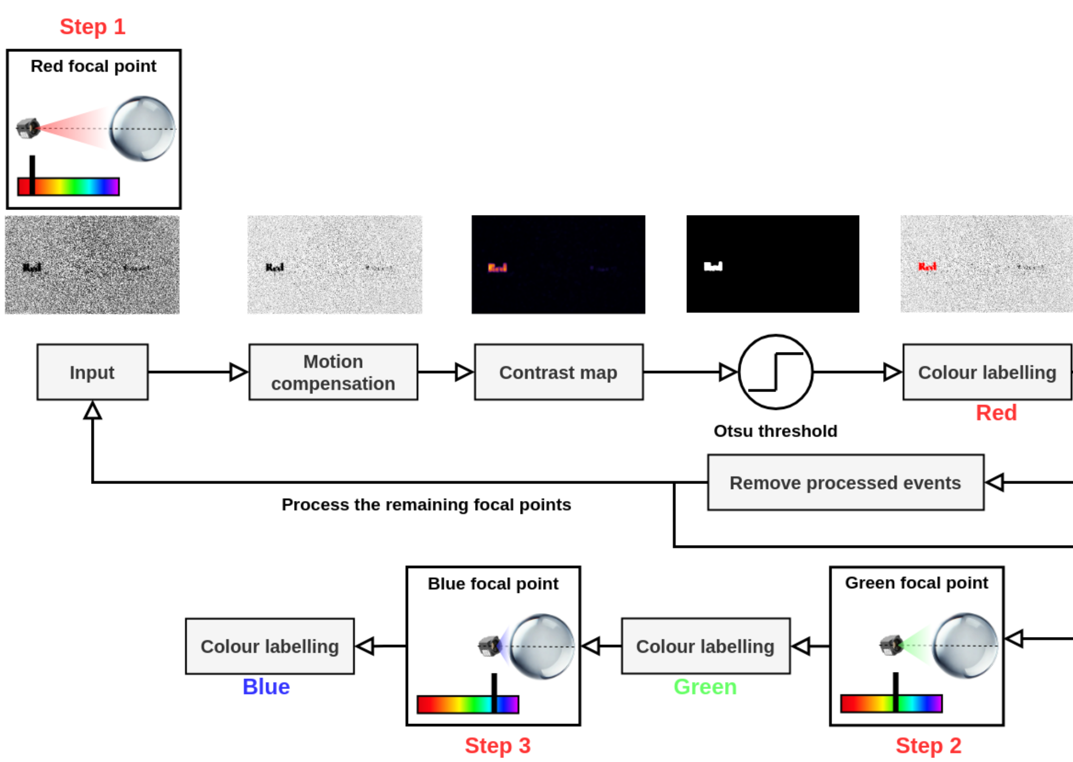

Fig. 3: Colour segmentation algorithm: A step-by-step procedure to automatically label event streams based on local sharpness in the accumulated event image, effectively mapping events to their corresponding spectral information.

We propose a simple algorithm (Figure 3) to segment spectral information from event camera data. Our dataset was captured in an uncontrolled environment, where a cellphone screen displayed "Red," "Green," and "Blue" in their respective colours. The algorithm assumes known focal distances for specific wavelengths and shifts the sensor's focal plane accordingly. When focused on a particular colour, events corresponding to that wavelength appear at distinct spatial locations.

Before running the segmentation, we perform a calibration step. Using a spectrometer, we identify the linear actuator positions that yield peak intensity (sharpest focus) for each wavelength. This process is repeated for all target wavelengths, resulting in an accurate actuator-to-wavelength mapping.

Once calibrated, the algorithm systematically moves the sensor back and forth through the defined focal ranges. During each sweep, it accumulates all triggered events into a single event image by counting events at each pixel. To enhance clarity, we then apply a Contrast Maximisation (CMax) algorithm [2], assuming linear motion, to effectively deblur and compensate for small movement.

Next, a 30×30 sliding window traverses over the accumulated image to measure local sharpness by calculating variance [1] within each window; higher variance indicates sharper focus. We binarise this variance map using Otsu's thresholding to isolate the sharpest regions. Events in these regions are labeled according to their spectral content, then removed from subsequent processing. The actuator moves to the next focal plane, and the process repeats until all wavelengths have been segmented. The number of colour labels depends on how many wavelengths to focal points are known during the calibration step. In future work, we aim to develop a technique that doesn't require a calibration step to perform spectral segmentation in a single shot.

We evaluated this approach on both static (only focal length is changing) and dynamic (focal length changing with multiple moving objects) scenes, shown in Figures 4 and 5. Our results confirm the viability of colour-from-focus in uncontrolled settings. Figure 4(a) illustrates how the camera-linear actuator setup can focus on one colour while blurring others, even when the displayed colours are not purely red, green, or blue. Although some overlap occurs (e.g., green events appearing in blurred red regions), the dominant signal remains sufficiently distinct for reliable detection. Adjusting the camera bias threshold or using a more precise actuator can further reduce overlap; this is left as future work. Figure 4(b) shows how the focused images at different wavelengths can be combined to form a superimposed RGB image.

Finally, Figure 5 demonstrates successful segmentation for dynamic scenes with randomly moving shapes. In these cases, events are not discarded after each focal plane step, since object and camera motion cause event distributions to shift over time. Our results confirm that accurate spectral segmentation is achievable in moving scenarios, provided the linear actuator can reliably adjust focal positions.

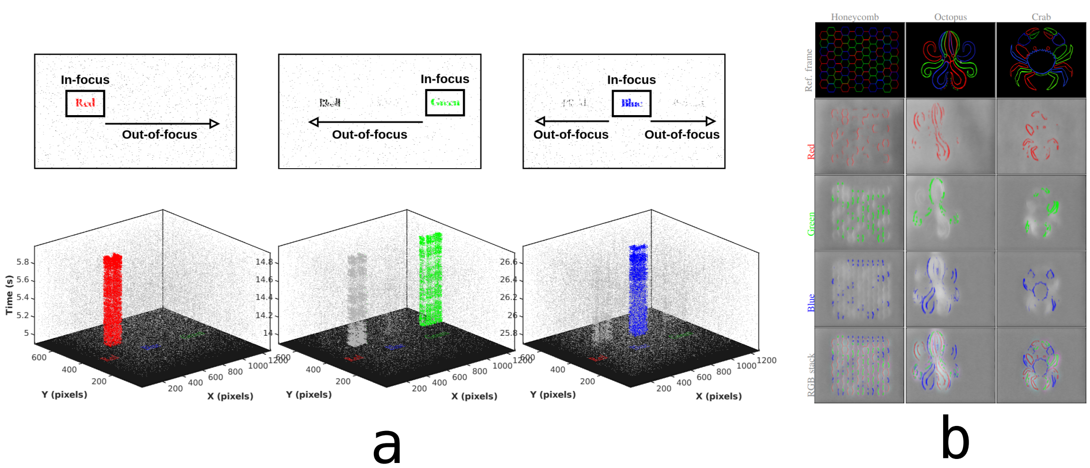

Fig. 4: Colour segmentation on 2D objects. (a) Segmentation of three objects that are separated from each other. (b) Segmentation of objects that are close to each other. Both show how the cephalopod-inspired event camera can perceive colours and how they are segmented.

References

[1] Gallego, G., Gehrig, M., Scaramuzza, D.: Focus Is All You Need: Loss Functions For Event-based Vision. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 12272–12281 (Jun 2019). https://doi.org/10.1109/CVPR.2019.01256, http://arxiv.org/abs/1904.07235, arXiv:1904.07235 [cs]

[2] Gallego, G., Rebecq, H., Scaramuzza, D.: A Unifying Contrast Maximization Framework for Event Cameras, with Applications to Motion, Depth, and Optical Flow Estimation. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3867–3876. IEEE, Salt Lake City, UT (Jun 2018). https://doi.org/10.1109/CVPR.2018.00407, https://ieeexplore.ieee.org/document/8578505/

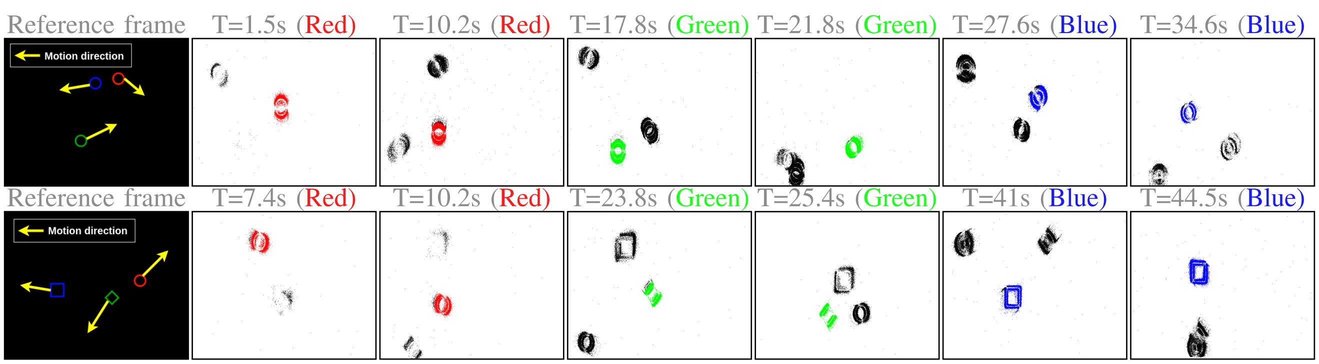

Fig. 5: Colour segmentation on coloured moving objects. Sharp and blurry objects trigger events, but the sharp produce more contrast and therefore the algorithm can segment them at each focal plane.