High-fidelity Event-Radiance Recovery via Transient Event Frequency

Jin Han[1,2] Yuta Asano[2] Boxin Shi[† 3,4] Yinqiang Zheng[1] Imari Sato[1,2]

[1]Graduate School of Information Science and Technology, The University of Tokyo [2]National Institute of Informatics

[3]National Key Laboratory for Multimedia Information Processing, School of Computer Science, Peking University

[4]National Engineering Research Center of Visual Technology, School of Computer Science, Peking University

{jinhan, asanoy, imarik}@nii.ac.jp shiboxin@pku.edu.cn yqzheng@ai.u-tokyo.ac.jp

Abstract

High-fidelity radiance recovery plays a crucial role in scene information reconstruction and understanding. Conventional cameras suffer from limited sensitivity in dynamic range, bit depth, and spectral response, etc. In this paper, we propose to use event cameras with bio-inspired silicon sensors, which are sensitive to radiance changes, to recover precise radiance values. We reveal that, under active lighting conditions, the transient frequency of event signals triggering linearly reflects the radiance value. We propose an innovative method to convert the high temporal resolution of event signals into precise radiance values. The precise radiance values yields several capabilities in image analysis. We demonstrate the feasibility of recovering radiance values solely from the transient event frequency (TEF) through multiple experiments.

1. Introduction

Scene radiance recovery from pixel values of conventional frame-based cameras is challenging due to the limited sensitivity of sensors and the non-linear process in the image signal processor. Event cameras like Dynamic Vision Sensor (DVS) [16] are designed with bio-inspired mechanism that measure the scene radiance changes in an asynchronous manner. Compared with conventional frame-based cameras, event cameras have superior advantages [7], such as very high temporal resolution (in the order of μs), high dynamic range (HDR, up to 120dB), low latency, low power consumption, *etc*. Due to the specific event triggering mechanism, event cameras only record radiance changes rather than the absolute radiance values. This makes it challenging to directly apply computer vision algorithms designed for 2D images containing luminance values for most of the surface points to the data captured by event cameras. We propose an innovative method for high-fidelity radiance recovery, which benefits various spectroscopic and stereoscopic vision fields, as shown in Fig. 1.

† Corresponding author

Project page: https://github.com/hjynwa/TEF

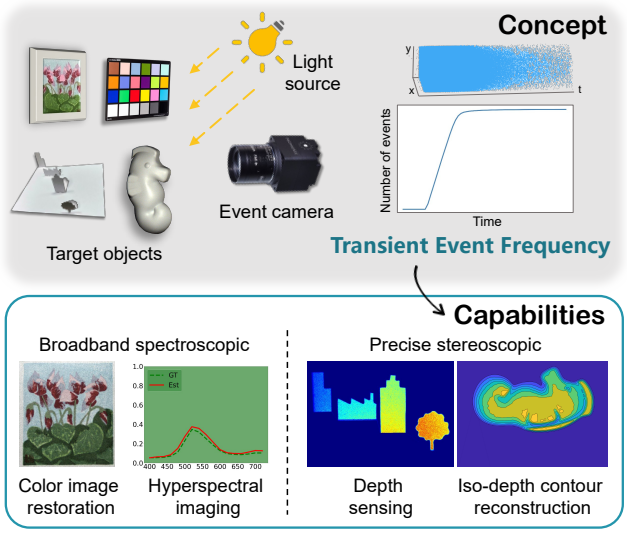

Figure 1. The concept of transient event frequency and its capabilities. When turning the light source on, an event camera captures event signals in a very short time, the frequency of event triggering during this period encodes the scene radiance. The capabilities of transient event frequency are experimentally verified by various applications, including color image restoration, hyperspectral imaging, depth sensing, and iso-depth contour reconstruction.

Since the event signals could be triggered from either lighting condition changes or object motions, there are many methods [14, 27, 29, 44] assuming a constant lighting condition to focus on motion. To acquire more *dense* event signals used for scene analysis, some researchers have proposed to use active lighting that can reflect the information about the entire scene from the event signals [20, 32]. Active lighting can be a transient state by turning on the light, which only takes less than 0.1 second. Under such a circumstance, the event signals are densely triggered for almost all surface points in the scene, rather than sparsely along the edges. Previous methods [6, 32] have shown that radiance can be recovered by integrating a period of events. However, this approach may suffer from unreliability due to the presence of ghost signals that persist even after the radiance change has ceased [21], resulting in errors in the recovered radiance values. This is caused by the latency and noise

of event triggering, which make it difficult to determine the precise termination timestamp for signal accumulation. As a result, integrating events can be challenging when there are extremely large radiance changes.

In this paper, a new approach for direct recovery of scene radiance from event signals is proposed. We overcome the instability and errors of event signals caused due to large luminance changes in a split second by analyzing the transient event frequency (TEF). Events are triggered in a constant frequency during the period of illumination increase, which linearly represents the radiance values of a point. As shown in Fig. 1, our approach transfers the high temporal resolution of event signals into the precise radiance values, and thus inherits the advantages of event cameras, which have low latency, broader spectral and dynamic range response than conventional cameras. Furthermore, we can directly apply computer vision algorithms developed for 2D images to event-radiance images reconstructed from event signals, expanding the potential applications of event cameras.

To the best of our knowledge, this study is the first to show that the high temporal resolution of event signals can be converted to relative radiance values by analyzing the triggering frequency under active lighting. Our contributions are summarized as follows:

1) We propose the concept of TEF during the split second of turning light on. It precisely represents the relative radiance values in a scene, which is much more stable and accurate compared to the integration of events and the pixel values from conventional cameras.

2) We reveal the linear relationship between TEF and radiance values, which yields several capabilities, including color image restoration, hyperspectral measurement, depth sensing, and iso-depth contour reconstruction.

3) We calibrate the linearity of TEF and measure the spectral response function of the event camera. Multiple experiments validate the broad response and robust precision of radiance values recovered from TEF in spectroscopic and stereoscopic vision fields.

2. Related Works

Event cameras are sensitive to radiance changes at high temporal resolution with HDR, which could be induced by objects movement or the illumination changes in active lighting.

2.1. Event-based vision triggered by motion

Under static lighting conditions, event streams are primarily triggered by scene and/or camera motion. Thanks to the advantages of event cameras, they are being imposed to computer vision and robotics fields to solve versatile tasks. In intensity reconstruction, event cameras have been applied to improve the quality of restored images and videos, such as image deblurring [5, 23, 31, 39, 41], optical flow estimation [1, 22, 30, 42], high-speed video reconstruction [18, 27, 29, 44], high-frame-rate (HFR) video synthesis [25, 33, 34], and HDR imaging [10, 11, 38, 40], *etc*. Besides, there are many works applying event cameras to stereoscopic vision tasks, like depth estimation [9, 12, 26], shape reconstruction [4, 37], SLAM [19, 24], *etc*. Since the events are mainly triggered in the edges of objects, event cameras can replace conventional cameras in many high-level vision tasks that utilize edge information, such as object segmentation [3, 35], tracking [2, 8], and pose recognition [36, 43]. These event-based methods rely on the events triggered by motion, assuming that the lighting conditions remain unchanged during the process of capturing.

2.2. Event-based vision under active lighting

Illumination changes are another source to trigger events of a scene. Research utilizing active lighting for event triggering can be categorized into two categories: structured light and intensity-changing light. The structured light with different patterns illuminates the target objects, whose reflection is captured by event cameras with high temporal resolution. MC3D [17] exploited the mapping between spatial disparity of a laser point projector and the temporal information encoded in event streams for 3D shape recovery. Huang *et al*. [13] used a high-speed digital light projector (DLP) as the light source, and proposed a digital image correlation method to calculate the displacements for 3D surface reconstruction. ESL [20] maximized the spatio-temporal correlation between the projector and a event camera to suppress the effect of noise in event streams. Besides structured light, researchers use light sources with intensity changes to trigger event signals. Takatani *et al*. [32] proposed to use an event camera for bispectral photometry with temporally modulated light. The light modulated at two individual wavelengths illuminated the target medium, while the triggered events recorded the bispectral difference, which can be used for depth estimation and turbid medium concentration estimation. Chen *et al*. [6] analyzed the event streams triggered in the split second when the light is turned on and solved the problem of intensity-distance ambiguity for indoor lighting estimation.

Both of these methods [6, 32] tried to recover radiance values by integrating a period of events. But we focus on the event triggering frequency during the transient illumination changes.

3. Preliminaries

3.1. Event signals triggering

An event camera detects the changes of radiance in a scene and outputs a sequence of event stream $\mathbb{E} = \{\mathbf{e}_i\}_{i=1}^{N}$.

The radiance I is firstly amplified into logarithm domain by the photoreceptor in an event camera. When the logarithmic amplified radiance change exceeds the contrast threshold θ, an event $\mathbf{e}_i = (x, y, p, t)$ will be triggered, which encodes the coordinate (x, y) of the pixel, the polarity $p \in \{-1, 1\}$ indicating the increase or decrease of radiance values, and the timestamp t when the event is triggered. The process can be formulated as:

$$\Delta I \begin{cases} \geq \theta, & p = +1 \\ \in [-\theta, \theta], & \text{none} \\ \leq -\theta, & p = -1 \end{cases} \tag{1}$$

where $\Delta I = \log I_t^{(x,y)} - \log I_{t-\Delta t}^{(x,y)}$ represents the radiance changes in the pixel (x, y) during the period of Δt.

3.2. Radiance from integrating event signals

Previous research [6, 32] show that intensity values can be recovered by integrating a period of events, which is represented as:

$$I_{t_1} = I_{t_0} \cdot \exp\left(\theta \int_{t_0}^{t_1} \mathbb{E} dt\right), \tag{2}$$

where I_{t_1} and I_{t_0} are intensity values at timestamp t_1 and t_0, respectively. When the light is turned on, the illumination increases dramatically in a split second, which triggers tons of events in a very short period. This is known to result in a smearing effect when detecting sudden changes of motion [21]. Similarly, in the case of turning on a light, there often exits the tailing effect due to the latency and noise of event triggering. As Fig. 2 (a) shows, there exists the tailing effect even after the illumination becomes stable, which makes it difficult to decide the termination timestamp of event counting.

4. Transient Event Frequency

Instead of reconstructing radiance values in an integration way, we turn to focus on the event triggering frequency during the period of turning light on, which builds a much more accurate and robust connection to radiance values.

4.1. Radiance from event frequency

Consider a simple case where a white board with diffuse reflection (supposed as 1) and a light source. When turning the light on, the radiance of the white board directly reflects the illumination increase of the light source, which is captured by an event camera. The intensity of illumination changes can be formulated as an exponential function $\exp(\Phi(t))$ [6, 15] with respect to time t shown in Fig. 2 (b). After the logarithm amplification in event camera, the illumination changes can be represented as:

$$\Delta I = \Phi(t) - \Phi(t - \Delta t). \tag{3}$$

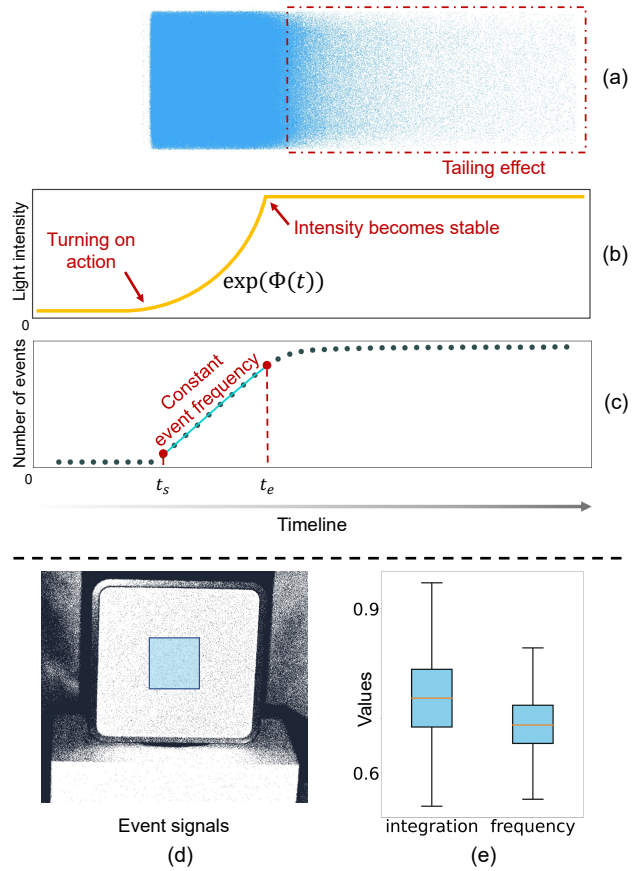

Figure 2. (a) A white board with diffuse reflection is illuminated by the light source, and triggers lots of event signals (the blue dots) with a tailing effect. (b) The intensity of illumination changes along the timeline. When turning on the light, the illumination increases dramatically and then becomes stable. The increase period can be formulated as an exponential function $\exp(\Phi(t))$. (c) The average number of events triggered in each pixel of the white board along the timeline. (d) A patch of pixels in the white board (blue masked) is selected for analysis. (e) The variance of the values from the integration-based method and our frequency-based method among the selected pixels.

According to Eq. (1), when ΔI exceeds the contrast threshold θ, an event signal will be triggered. Therefore, we can plot the number of events along the timeline to find its relationship to illumination increase in the split second of turning light on. As shown in Fig. 2 (c), after the turning on action, the average number of events increases linearly from t_s to t_e. It means that the event signals are triggered in a constant frequency especially in the central period of illumination increase, which is defined as the transient event frequency.

Therefore, $\Phi(t)$ can be easily approximated by a linear function[1]:

$$\Phi(t) = kt + b, \tag{4}$$

[1] What we approximate is just the linearly increasing part of the number of events (*i.e.*, t_s to t_e in Fig. 2 (c)), not the whole curve.

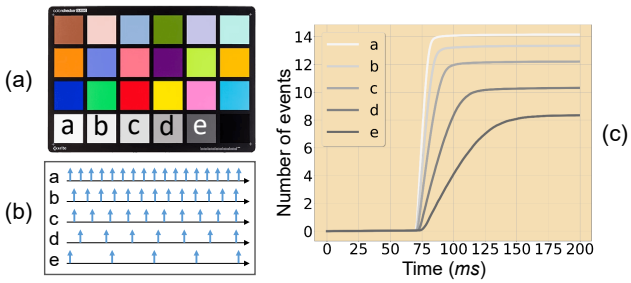

Figure 3. (a) The first 5 patches in the last row of a Macbeth ColorChecker are selected for comparison. (b) The events triggered in different patches along timeline, where blue up-arrows represent the positive ("+1") events. (c) The number of events increasing during the split second of turning light on.

where k is the slope factor and b is the constant term. The slope factor k reflects different radiance values reflected from the scene. As shown in Fig. 3, when turning on the light source, the amount of radiance increase of the white patch (labeled as "a") in a Macbeth ColorChecker is larger than that from the gray patch (labeled as "e"), both from the total dark. Therefore, patches with higher radiance values trigger events in a higher frequency, as shown in Fig. 3 (b). We plot the number of events along timeline. The larger slope factor k, or a higher TEF represents a higher radiance value.

During the period of radiance increase (*e.g.*, from t_s to t_e), a sequence of event signals $\{\mathbf{e}_i\}_{i=1}^{N}$ will be triggered in each pixel. The time interval Δt_i between two consecutive events is computed by $t_{i+1} - t_i$. Therefore, the TEF at pixel (x, y) can be formulated as:

$$\begin{aligned} f^{(x,y)} &= \frac{N-1}{\sum_{i=1}^{N-1} \Delta t_i^{(x,y)}} \\ &= \frac{1}{\Delta \bar{t}^{(x,y)}}, \end{aligned} \tag{5}$$

where $\Delta \bar{t}^{(x,y)}$ is the averaged timestamp interval of event signals triggered at (x, y) during the period of illumination increase (from t_s to t_e).

According to the definition of events integration, all the events triggered *before* the light becoming constant should be counted for computing the changes of radiance. However, for TEF, the ending time bounds (t_e) are earlier because we intercept the period when events are triggered in a constant frequency[2]. In fact, TEF can be computed from any range within $[t_s, t_e]$, which is an advantage over integration of events. Averaging more time intervals can increase the robustness of TEF. As shown in Fig. 2 (d), we select a patch of pixels from the white board, and compute their values from integration-based method and our frequency-based method. The values of different pixels should be identical ideally. Figure 2 (e) shows that the variance of values computed from the integration is much larger than that from our frequency-based method, which demonstrates that our frequency-based method is much more stable and robust. The superiority is more obvious in low-light conditions because of the difficulty in determining the ending time bound for integration.

[2]Details of how to determine the time bounds for frequency and integration are described in the supplementary material.

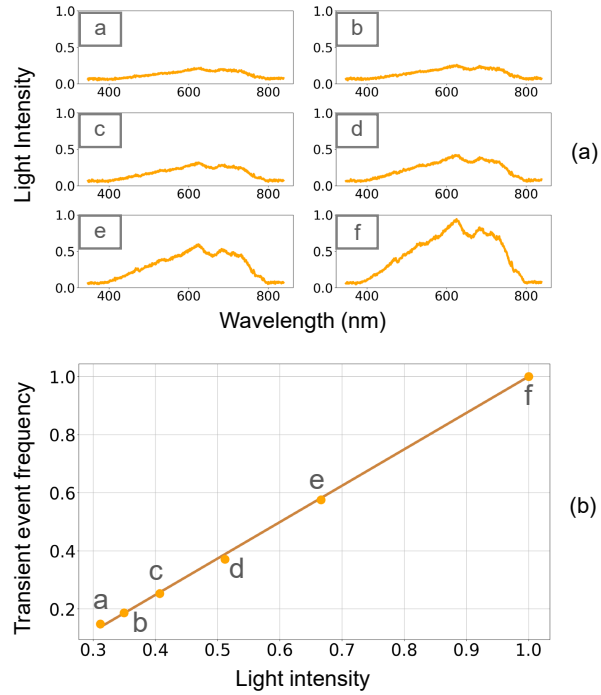

Figure 4. (a) The intensities of the light source at 6 different distances from the white board. (b) The relationship between TEF and light intensity, which can be fitted by a linear curve.

4.2. Linearity under active lighting

To verify the linearity of TEF with respect to radiance values, we need to change the intensity of lighting and record TEF under different lighting. We vary the intensity of the light source at 6 different distances (in the same direction) from the white board, and compute the TEF of each point. The light source intensities are measured using a spectrometer by integrating the intensities across all wavelengths (from 380nm to 800nm). As shown in Fig. 4, we plot the TEF values for the 6 points (all the values are normalized to $[0, 1]$) and find that they could be easily fit using a linear curve, which demonstrates the linearity of TEF with respect to radiance values.

4.3. Measuring event spectral response

Measuring the spectral response curve of event cameras is essential for acquiring accurate color information from

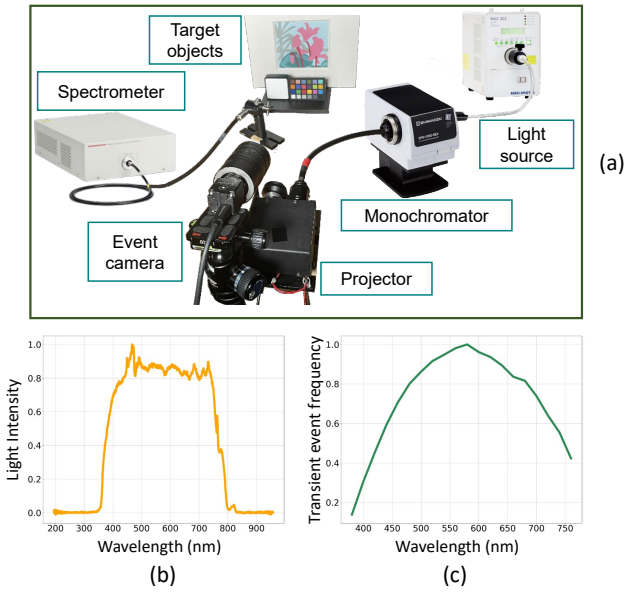

Figure 5. (a) The experimental setup to obtain the spectral response function of an event camera. (b) The spectral distribution of the light source. (c) The spectral response curve of a Prophesee EVK4 event camera that we measured in the range of 380nm to 760nm.

event signals. For event spectral response measurement, we use a monochromator and a spectrometer to scan over the wavelength range of interest (from 380nm to 760nm with an interval of 10nm). A stable light source with a smooth spectral distribution and sufficient intensity among the wavelength is used. Besides, we use a projector to act as a switch that controls turning light on and off. As shown in Fig. 5 (a), we use a white board with uniform reflectivity across all wavelengths as the calibration object. The monochromator produces narrow-band lights, while the spectrometer measures the intensity of each narrow-band light.

Since we have demonstrated the linearity of TEF in Sec. 4.2, the spectral response function of event cameras can be computed by $C(\lambda) = L(\lambda)/f(\lambda)$, where $f(\lambda)$ is the TEF under illumination $L(\lambda)$. We plot the spectral response curve of the event camera that we use (Prophesee EVK4) in Fig. 5 (c).

4.4. Scene analysis

The high-fidelity radiance values recovered via TEF benefit several fields, including hyperspectral (multi-band) measurement, depth sensing, and iso-depth contour reconstruction.

Hyperspectral (Multi-band) measurement. Since, the TEF builds a linear relationship with the scene radiance, we can measure the spectral reflectance $R(\lambda)$ of a scene using narrow-band illuminants. The intensity value of a pixel in

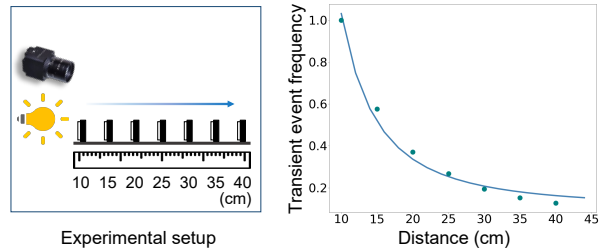

Figure 6. Left: The experimental setup for validating the relationship between distance and TEF. Right: We plot the 7 data points and fit them using a reciprocal of quadratic function described in Eq. (10).

an image is modeled as the product of three factors: the illumination $L_n(\lambda)$, the camera's spectral response function $C(\lambda)$, and the spectral reflectance of this pixel $R^{(x,y)}(\lambda)$. Since the TEF has a linear response to scene radiance $I_n^{(x,y)}$, the TEF can be formulated as:

$$f_n^{(x,y)} \propto I_n^{(x,y)} = \int L_n(\lambda)C(\lambda)R^{(x,y)}(\lambda)d\lambda, \quad (6)$$

where λ is the wavelength, $f_n^{(x,y)}$ is the TEF at pixel (x, y) under the n-th illumination. Since we use a monochromator to produce narrow-band illuminants, Eq. (6) can be rewritten as:

$$f_n^{(x,y)} = L_n(\lambda_n)C(\lambda_n)R^{(x,y)}(\lambda_n), \quad (7)$$

where λ_n is the peak wavelength of the n-th narrow-band illuminant. Then, the spectral reflectance of wavelength λ_n at pixel (x, y) can be computed by:

$$R^{(x,y)}(\lambda_n) = \frac{f_n^{(x,y)}}{L_n(\lambda_n)C(\lambda_n)}. \quad (8)$$

Depth sensing. The relationship between light intensity and distance can be formulated according to the inverse square law of light fall-off property:

$$L_d = L \cdot \frac{1}{4\pi d^2}, \quad (9)$$

where L_d is the light intensity at the spherical surface with a radius d, and the light source L is the center of the sphere. So L_d is proportional to $\frac{1}{d^2}$ (*e.g.*, $L_d \propto \frac{1}{d^2}$). When we increase the distance of target objects from the light source, the radiance reflected from the target objects decreases. The TEF is sensitive to small differences in radiance, which directly reflects the depth values. We use a white board with diffuse reflection as the target object, and increase the distance between the white board and the light source linearly from 10cm to 40cm with an interval of 5cm. As shown in Fig. 6, the relationship between distance and TEF can be fitted by a reciprocal of quadratic function:

$$f_d = a \cdot \frac{1}{d^2} + b, \quad (10)$$

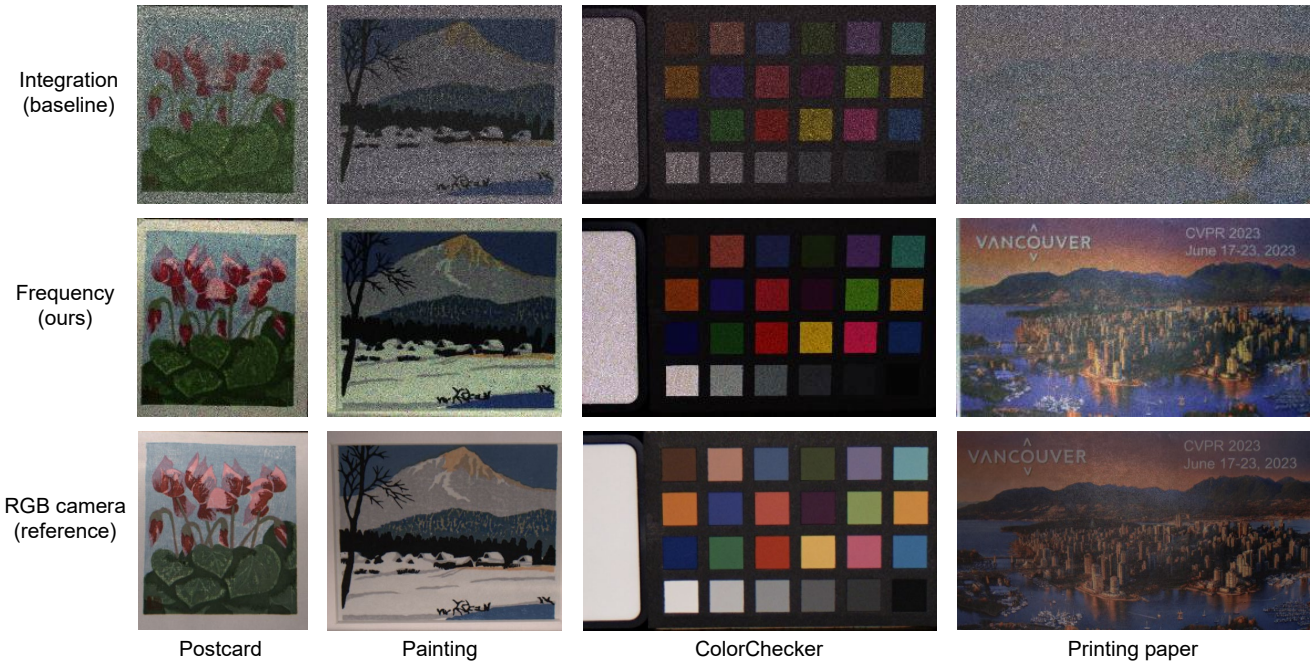

Figure 7. The results of color image restoration. The first row is the results from integrated events using Eq. (2). The second row is the results using the proposed TEF. The final row is the images captured with an RGB camera.

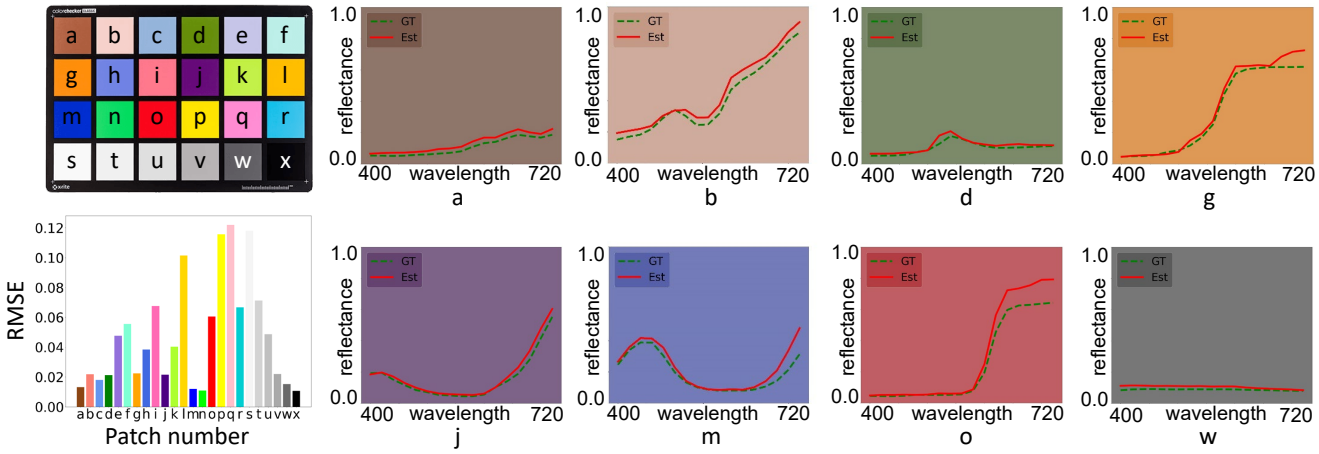

Figure 8. A Macbeth ColorChecker is used to evaluate the accuracy of reconstructed spectral reflectance curves. 8 patches are selected as samples to show the similarity between our estimation (red solid line labeled as "Est") and the ground truth (green dotted line labeled as "GT"). The RMSE of all patches are plotted in the bottom left chart.

where f_d is the TEF at distance d. It directly corresponds to the inverse square law of light fall-off property in Eq. (10). So the depth information can be easily distinguished by the TEF values.

Iso-depth contour reconstruction. Assuming there are parallel rays of distant light and a Lambertian surface with uniform albedo, the radiance value of a point in the surface can be formulated as:

$$R^{(x,y)} = \rho L \cos(\alpha), \tag{11}$$

where ρ is the albedo, L is the light intensity, and α is the angle between the light direction and the surface normal of point (x, y). The angle α decides the radiance value of a point in the surface (*e.g.*, larger α leads to lower radiance value). Since TEF is linearly related to radiance value, it is possible to estimate a conical surface containing the surface normal of each point in the surface. The points with surface normals in the same conical surface can be clustered into the same contour line. This allows us to reconstruct the iso-depth contours of the differentiable surface of an object.

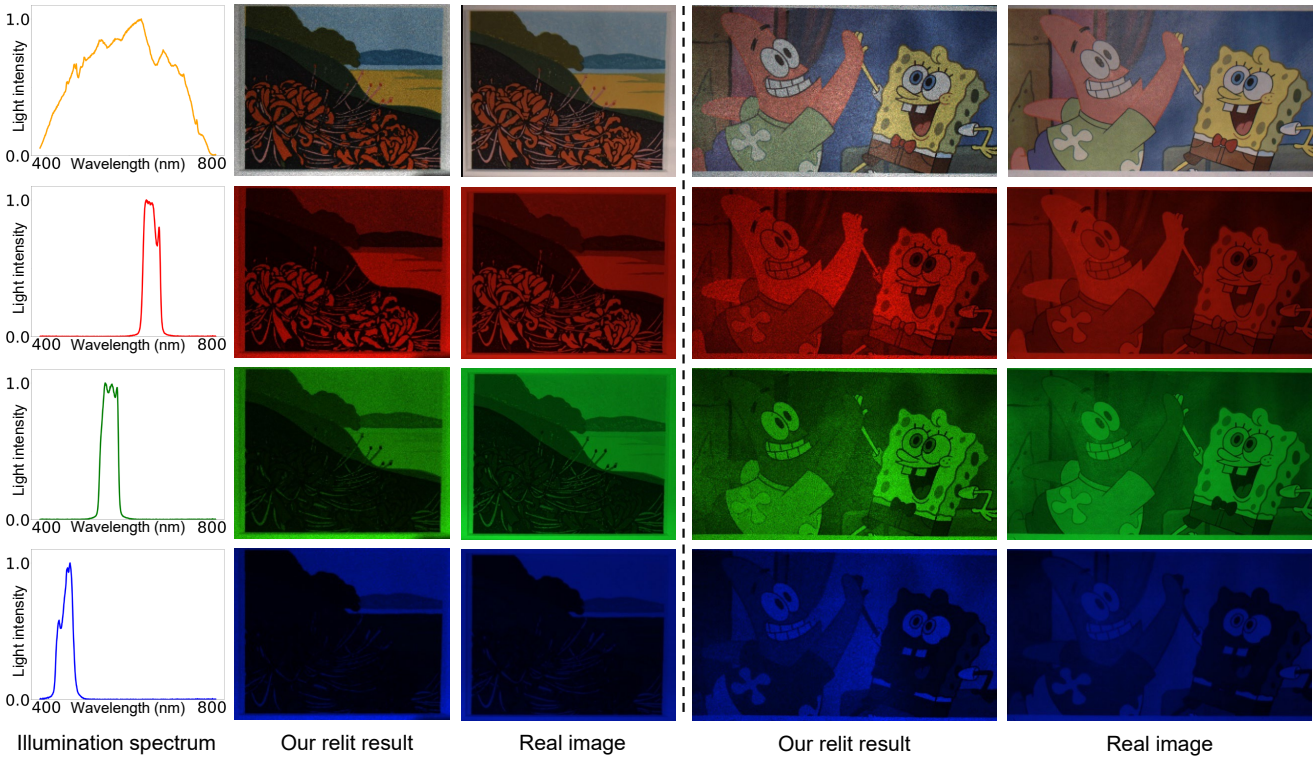

Figure 9. Image relighting results computed from the recovered spectral reflectance curves under 4 different illuminations. The illumination spectrums are measured by a spectrometer. The real images are captured using an RGB camera.

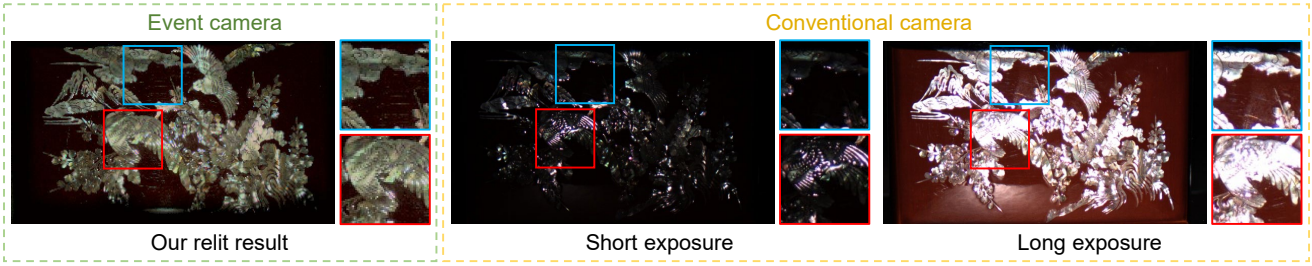

Figure 10. An HDR example of image relighting for specular surface.

5. Experiments

5.1. Color image restoration

Since event cameras only measure the radiance changes, no color information is captured. We apply red, green, and blue filters in front of the light source and capture the TEF 3 times for R, G, and B channel, respectively. The pixel values of color images restored from transient frequency can be viewed as a linear transformation from scene radiance values. As shown in Fig. 7, we compare the color images restored from TEF (ours) and the integration-based method (baseline) used in [6, 32]. The pictures in the final row are captured using a conventional RGB camera for reference. The images restored from TEF and integration baseline have the same input event signals. Our frequency-based results are less affected by the noise, which have more detailed color appearance. The results demonstrate that the TEF can restore scene radiance values with higher fidelity compared to integration of events.

5.2. Hyperspectral (Multi-band) imaging

We reconstruct the spectral reflectance at each patch[3] using the monochromator scanning from 380nm to 760nm with an interval of 20nm. To evaluate spectral accuracy of our reconstruction, a 24-patch Macbeth ColorChecker with ground truth spectral reflectance curves is used as the standard object. As shown in Fig. 8, the spectral reflectance curves reconstructed with TEF are similar to the ground truth curves. Besides, we compute the root mean square

[3] We consider each 2×2 pixels as a small patch to suppress noise.

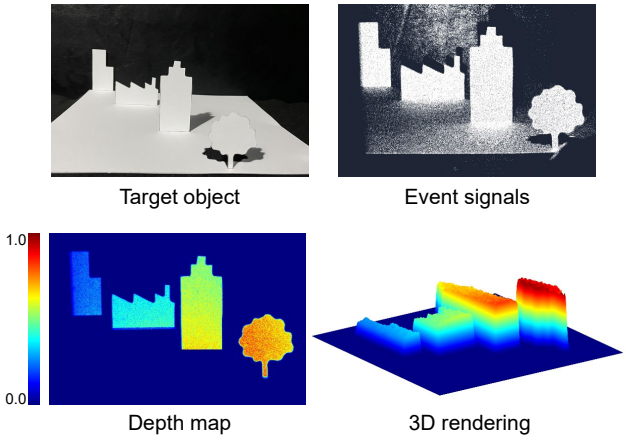

Figure 11. Depth sensing result. Top left: The target object that is made of diffuse paper. Top right: The visualized event signals triggered during the period of all wavelengthsing light on. Bottom left: The reconstructed depth map. Bottom right: A 3D rendering of the object.

error (RMSE) of the 24 patches. The biggest error coming from the pink patch (labeled as "q") is smaller than 0.125. The results demonstrate the accuracy of spectral reflectance recovery via TEF both qualitatively and quantitatively.

Given the recovered spectral reflectance curves of the scene, we can conduct spectral relighting using Eq. (6) with known illumination spectrum. As shown in Fig. 9, the relit results appear very similar to real images captured by a conventional RGB camera (FLIR Grasshopper GS3-U3-28S5C-C) under different illuminants. Thanks to the HDR property, event cameras can preserve texture details of a scene with high dynamic range, while conventional cameras suffer from saturation or noise effect due to the limited dynamic range of their sensors. An example of a box lid with specular textures is shown in Fig. 10. We firstly measure the spectral reflectance of the surface using TEF. Then we relight the image under the illumination shown in the first column in Fig. 9. The HDR result is tone mapped using the method [28] for visualization. It is obvious that the image reconstructed from TEF restores both high-radiance specular and low-radiance parts of the surface accurately.

5.3. Depth sensing

We design an object with multiple paper layers lying at different levels and capture the events triggered during the split second of turning light on. Then we measure the transient event frequencies of different paper layers. Since all the layers are faced to the same direction, their radiance values are only affected by the distance to the light source. As shown in Fig. 11, the event camera can successfully distinguish different layers based on the values of TEF. By creating a look-up table according to Fig. 6, under the same lighting setup, we can even recover the actual depth values of the target objects. Such differences of depth values are hard to distinguish by conventional cameras.

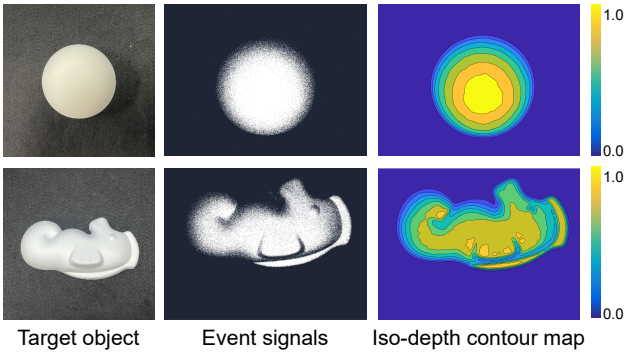

Figure 12. Iso-depth contour reconstruction results. Left column: The target objects. Middle column: The visualized event signals triggered during the period of turning light on. Right column: The reconstructed iso-depth contour maps.

5.4. Iso-depth contour reconstruction

We use an event camera to capture several target objects with a diffuse surface, and compute the TEF at each small patch. The values of frequency represent the angles between the surface normals and the direction of the light source. Therefore, we can plot the iso-depth contour map by clustering the points that have the same TEF values. The results are shown in Fig. 12.

6. Conclusion

In this paper, we propose the concept of transient event frequency, which directly reflects the scene radiance. We can recover the precise scene radiance values by converting the high temporal resolution of event signals during the split second of turning light on into radiance value differences. TEF yields various capabilities in spectroscopic and stereoscopic vision, as demonstrated by several experimental applications.

Limitations and future work. Although we can recover the precise radiance values by analysing the event triggering frequency, our method requires a specific setup, *e.g.*, an action of turning on a light, which cannot be directly applied to outdoor scenarios with natural lighting. Besides, we assume the event camera and objects in the scene are both static, as motion may also trigger event signals. Separating event signals triggered from illumination changes and movement is left for future work.

Acknowledgement

This work was supported by the National Key R&D Program of China under Grant No. 2021ZD0109803, the National Natural Science Foundation of China under Grant No. 62088102, 62136001, and the JSPS KAKENHI under Grant No. JP20H05953, JP22H00529.

References

[1] Patrick Bardow, Andrew J Davison, and Stefan Leutenegger. Simultaneous optical flow and intensity estimation from an event camera. In *Proc. of Computer Vision and Pattern Recognition*, 2016. 2

[2] Francisco Barranco, Cornelia Fermuller, and Eduardo Ros. Real-time clustering and multi-target tracking using event-based sensors. In *Proc. of International Conference on Intelligent Robots and Systems*, 2018. 2

[3] Francisco Barranco, Ching L Teo, Cornelia Fermuller, and Yiannis Aloimonos. Contour detection and characterization for asynchronous event sensors. In *Proc. of Internatoinal Conference on Computer Vision*, 2015. 2

[4] Kenneth Chaney, Alex Zihao Zhu, and Kostas Daniilidis. Learning event-based height from plane and parallax. In *Proc. of International Conference on Intelligent Robots and Systems*, 2019. 2

[5] Haoyu Chen, Minggui Teng, Boxin Shi, Yizhou Wang, and Tiejun Huang. A residual learning approach to deblur and generate high frame rate video with an event camera. *IEEE Transactions on Multimedia*, 2022. 2

[6] Zehao Chen, Qian Zheng, Peisong Niu, Huajin Tang, and Gang Pan. Indoor lighting estimation using an event camera. In *Proc. of Computer Vision and Pattern Recognition*, 2021. 1, 2, 3, 7

[7] Guillermo Gallego, Tobi Delbrück, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew J Davison, Jörg Conradt, Kostas Daniilidis, et al. Event-based vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020. 1

[8] Daniel Gehrig, Henri Rebecq, Guillermo Gallego, and Davide Scaramuzza. Asynchronous, photometric feature tracking using events and frames. In *Proc. of European Conference on Computer Vision*, 2018. 2

[9] Daniel Gehrig, Michelle Rüegg, Mathias Gehrig, Javier Hidalgo-Carrió, and Davide Scaramuzza. Combining events and frames using recurrent asynchronous multimodal networks for monocular depth prediction. *IEEE Robotics and Automation Letters*, 2021. 2

[10] Jin Han, Yixin Yang, Peiqi Duan, Chu Zhou, Lei Ma, Chao Xu, Tiejun Huang, Imari Sato, and Boxin Shi. Hybrid high dynamic range imaging fusing neuromorphic and conventional images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023. 2

[11] Jin Han, Chu Zhou, Peiqi Duan, Yehui Tang, Chang Xu, Chao Xu, Tiejun Huang, and Boxin Shi. Neuromorphic camera guided high dynamic range imaging. In *Proc. of Computer Vision and Pattern Recognition*, 2020. 2

[12] Javier Hidalgo-Carrió, Daniel Gehrig, and Davide Scaramuzza. Learning monocular dense depth from events. In *Proc. of International Conference on 3D Vision*, 2020. 2

[13] Xueyan Huang, Yueyi Zhang, and Zhiwei Xiong. High-speed structured light based 3D scanning using an event camera. *Optics Express*, 2021. 2

[14] Hanme Kim, Ankur Handa, Ryad Benosman, Sio-Hoi Ieng, and Andrew Davison. Simultaneous mosaicing and tracking with an event camera. In *Proc. of the British Machine Vision Conference*, 2014. 1

[15] Hyun-Ho Kim, Sang-Hyun Choi, Sang-Hyun Shin, Young-Ki Lee, Seok-Moon Choi, and Sung Yi. Thermal transient characteristics of die attach in high power led pkg. *Microelectronics Reliability*, 2008. 3

[16] Patrick Lichtsteiner, Christoph Posch, and Tobi Delbruck. A 128×128 120 db $15\mu s$ latency asynchronous temporal contrast vision sensor. *IEEE Journal of Solid-State Circuits*, 2008. 1

[17] Nathan Matsuda, Oliver Cossairt, and Mohit Gupta. MC3D: Motion contrast 3D scanning. In *Proc. of Internatoinal Conference on Computational Photography*, 2015. 2

[18] Mohammad Mostafavi, Lin Wang, and Kuk-Jin Yoon. Learning to reconstruct hdr images from events, with applications to depth and flow prediction. *International Journal of Computer Vision*, 2021. 2

[19] Elias Mueggler, Basil Huber, and Davide Scaramuzza. Event-based, 6-dof pose tracking for high-speed maneuvers. In *Proc. of International Conference on Intelligent Robots and Systems*, 2014. 2

[20] Manasi Muglikar, Guillermo Gallego, and Davide Scaramuzza. ESL: Event-based structured light. In *Proc. of International Conference on 3D Vision*, 2021. 1, 2

[21] Matthew Ng, Zi Min Er, Gim Song Soh, and Shaohui Foong. Aggregation functions for simultaneous attitude and image estimation with event cameras at high angular rates. *IEEE Robotics and Automation Letters*, 2022. 1, 3

[22] Liyuan Pan, Miaomiao Liu, and Richard Hartley. Single image optical flow estimation with an event camera. In *Proc. of Computer Vision and Pattern Recognition*, 2020. 2

[23] Liyuan Pan, Cedric Scheerlinck, Xin Yu, Richard Hartley, Miaomiao Liu, and Yuchao Dai. Bringing a blurry frame alive at high frame-rate with an event camera. In *Proc. of Computer Vision and Pattern Recognition*, 2019. 2

[24] Xin Peng, Ling Gao, Yifu Wang, and Laurent Kneip. Globally-optimal contrast maximisation for event cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021. 2

[25] Chiara Plizzari, Mirco Planamente, Gabriele Goletto, Marco Cannici, Emanuele Gusso, Matteo Matteucci, and Barbara Caputo. E2 (GO) MOTION: Motion augmented event stream for egocentric action recognition. In *Proc. of Computer Vision and Pattern Recognition*, 2022. 2

[26] Henri Rebecq, Guillermo Gallego, Elias Mueggler, and Davide Scaramuzza. EMVS: Event-based multi-view stereo-3D reconstruction with an event camera in real-time. *International Journal of Computer Vision*, 2018. 2

[27] Henri Rebecq, René Ranftl, Vladlen Koltun, and Davide Scaramuzza. Events-to-video: Bringing modern computer vision to event cameras. In *Proc. of Computer Vision and Pattern Recognition*, 2019. 1, 2

[28] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda. Photographic tone reproduction for digital images. In *Proceedings of Conference on Computer Graphics and Interactive Techniques*, 2002. 8

[29] Cedric Scheerlinck, Nick Barnes, and Robert Mahony. Continuous-time intensity estimation using event cameras. In *Proc. of Asian Conference on Computer Vision*, 2018. 1, 2

[30] Shintaro Shiba, Yoshimitsu Aoki, and Guillermo Gallego. Secrets of event-based optical flow. In *Proc. of European Conference on Computer Vision*, 2022. 2

[31] Chen Song, Qixing Huang, and Chandrajit Bajaj. E-CIR: Event-enhanced continuous intensity recovery. In *Proc. of Computer Vision and Pattern Recognition*, 2022. 2

[32] Tsuyoshi Takatani, Yuzuha Ito, Ayaka Ebisu, Yinqiang Zheng, and Takahito Aoto. Event-based bispectral photometry using temporally modulated illumination. In *Proc. of Computer Vision and Pattern Recognition*, 2021. 1, 2, 3, 7

[33] Stepan Tulyakov, Alfredo Bochicchio, Daniel Gehrig, Stamatios Georgoulis, Yuanyou Li, and Davide Scaramuzza. Time Lens++: Event-based frame interpolation with parametric non-linear flow and multi-scale fusion. In *Proc. of Computer Vision and Pattern Recognition*, 2022. 2

[34] Stepan Tulyakov, Daniel Gehrig, Stamatios Georgoulis, Julius Erbach, Mathias Gehrig, Yuanyou Li, and Davide Scaramuzza. Time Lens: Event-based video frame interpolation. In *Proc. of Computer Vision and Pattern Recognition*, 2021. 2

[35] Lin Wang, Yujeong Chae, Sung-Hoon Yoon, Tae-Kyun Kim, and Kuk-Jin Yoon. Evdistill: Asynchronous events to end-task learning via bidirectional reconstruction-guided cross-modal knowledge distillation. In *Proc. of Computer Vision and Pattern Recognition*, 2021. 2

[36] Yanxiang Wang, Bowen Du, Yiran Shen, Kai Wu, Guangrong Zhao, Jianguo Sun, and Hongkai Wen. EV-gait: Event-based robust gait recognition using dynamic vision sensors. In *Proc. of Computer Vision and Pattern Recognition*, 2019. 2

[37] Ziyun Wang, Kenneth Chaney, and Kostas Daniilidis. EvAC3D: From event-based apparent contours to 3d models via continuous visual hulls. In *Proc. of European Conference on Computer Vision*, 2022. 2

[38] Ziwei Wang, Yonhon Ng, Cedric Scheerlinck, and Robert Mahony. An asynchronous kalman filter for hybrid event cameras. In *Proc. of Internatoinal Conference on Computer Vision*, 2021. 2

[39] Fang Xu, Lei Yu, Bishan Wang, Wen Yang, Gui-Song Xia, Xu Jia, Zhendong Qiao, and Jianzhuang Liu. Motion deblurring with real events. In *Proc. of Internatoinal Conference on Computer Vision*, 2021. 2

[40] Yixin Yang, Jin Han, Jinxiu Liang, Imari Sato, , and Boxin Shi. Learning event guided high dynamic range video reconstruction. In *Proc. of Computer Vision and Pattern Recognition*, 2023. 2

[41] Chu Zhou, Minggui Teng, Jin Han, Jinxiu Liang, Chao Xu, Gang Cao, and Boxin Shi. Deblurring low-light images with events. *International Journal of Computer Vision*, 2023. 2

[42] Alex Zihao Zhu, Liangzhe Yuan, Kenneth Chaney, and Kostas Daniilidis. Unsupervised event-based optical flow using motion compensation. In *Proc. of European Conference on Computer Vision*, 2018. 2

[43] Shihao Zou, Chuan Guo, Xinxin Zuo, Sen Wang, Pengyu Wang, Xiaoqin Hu, Shoushun Chen, Minglun Gong, and Li Cheng. EventHPE: Event-based 3D human pose and shape estimation. In *Proc. of Internatoinal Conference on Computer Vision*, 2021. 2

[44] Yunhao Zou, Yinqiang Zheng, Tsuyoshi Takatani, and Ying Fu. Learning to reconstruct high speed and high dynamic range videos from events. In *Proc. of Computer Vision and Pattern Recognition*, 2021. 1, 2

High-fidelity Event-Radiance Recovery via Transient Event Frequency – Supplementary Material –

Jin Han[1,2] Yuta Asano[2] Boxin Shi[† 3,4] Yinqiang Zheng[1] Imari Sato[1,2]

[1]Graduate School of Information Science and Technology, The University of Tokyo [2]National Institute of Informatics

[3]National Key Laboratory for Multimedia Information Processing, School of Computer Science, Peking University

[4]National Engineering Research Center of Visual Technology, School of Computer Science, Peking University

{jinhan, asanoy, imarik}@nii.ac.jp shiboxin@pku.edu.cn yqzheng@ai.u-tokyo.ac.jp

7. Time Bounds for Frequency and Integration

7.1. Time bounds comparison

The time bounds selection is a key difference between the integration-based method and our proposed frequency-based method. As shown in Fig. 13, the starting time bounds (t_s, labeled with blue star) are identical for both methods. But the ending time bounds for frequency (t_e^{freq}, labeled with red star) and integration (t_e^{inte}, labeled with green star) are different according to their definitions. For event integration, radiance values are computed based on the number of triggered events. Therefore, so we count the number of events until the light becomes constant. In contrast, for event frequency, since we focus on the period when events are triggered in a constant frequency, the ending time bound is earlier than that of integration. Take patch "a" and "e" in the ColorChecker for comparison, we can find that t_e^{inte} is much later than t_e^{freq} in the darker patch (*i.e.*, patch "e"), which will introduce more noise for integration of events. By intercepting the period with constant TEF, the radiance values recovered from frequency of events are more robust and stable, especially in darker area of a scene or low-light conditions.

7.2. Time bounds detection

The starting time bounds t_s and ending time bounds t_e^{freq} can be easily detected by checking the time interval. Specifically, for t_s, when the time interval between two consecutive events is smaller than a pre-defined threshold, it indicates that the light is turned on at this time stamp. For t_e^{freq}, when the time interval changes over the pre-defined threshold, it means that the lighting becomes constant. The thresholds can be chosen based on the characteristics of the event camera and the lighting conditions. We can change the thresholds to adjust the range intercepted for constant event frequency by increasing or decreasing the thresholds.

† Corresponding author

Project page: https://github.com/hjynwa/TEF

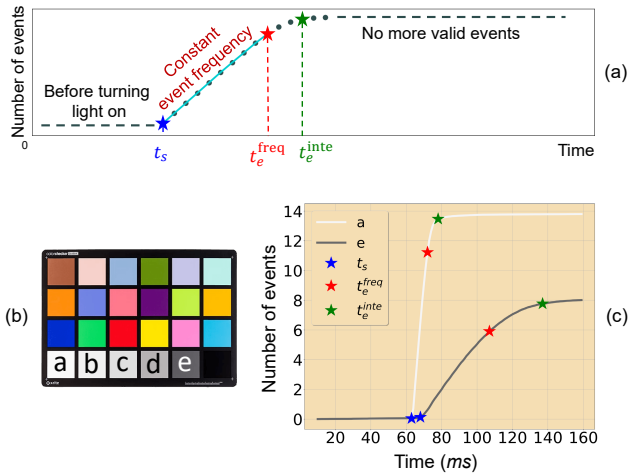

Figure 13. (a) The number of events changes along time when turning on a light. There exists a period when events are triggered in a constant frequency. (b) The MacBeth ColorChecker. (c) The curves of patch "a" and "e" labeled with t_s, t_e^{freq}, and t_e^{inte} are selected for comparison.

Please refer to our project page for implementation details.

8. More Experimental Results

We provide more experimental analysis and results to demonstrate the fidelity of the radiance recovered from the proposed TEF.

8.1. Hyperspectral imaging

By applying the narrow-band light spanning the wavelengths of interest, we can reconstruct the spectral curve of each point in a painting. The ground truth of spectral curves can be measured by a spectrometer with an fiber optical detector. As shown in Fig. 14, we measure the ground truth of spectral curve of a point, and compare it with our frequency-based estimation. In this figure, the spectral curves of the points indicated by yellow boxes are similar to the ground truth curves. We also provide additional image relighting

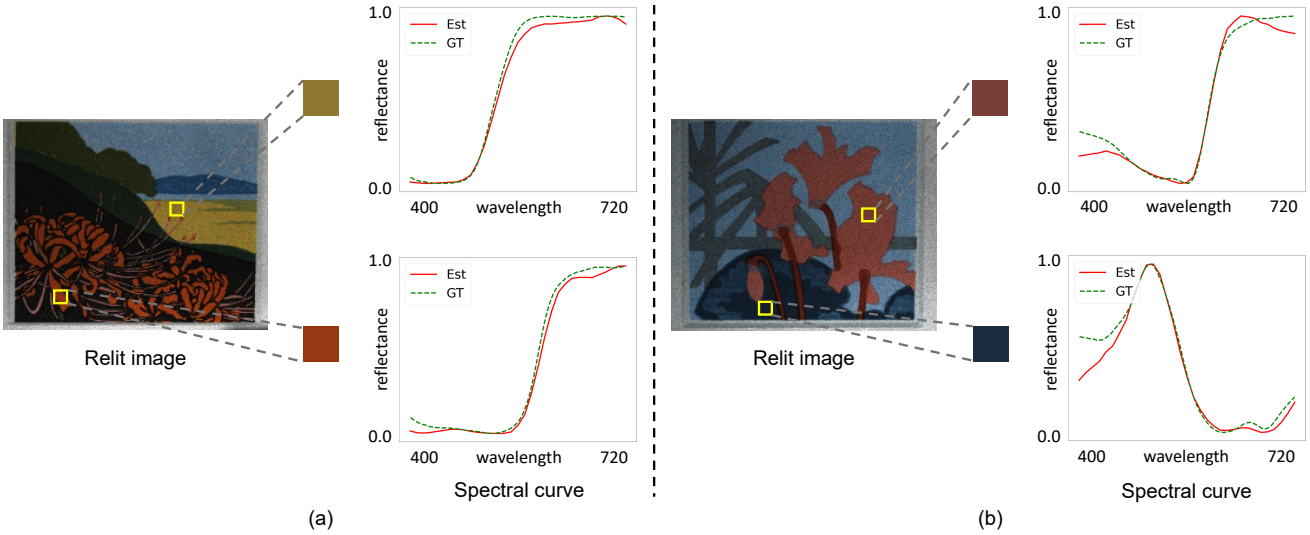

Figure 14. Two examples of spectral curve reconstruction. The points selected are enlarged and labeled with yellow boxes. The curves show the similarity between our estimation (red solid line labeled as "Est") and the ground truth (green dotted line labeled as "GT").

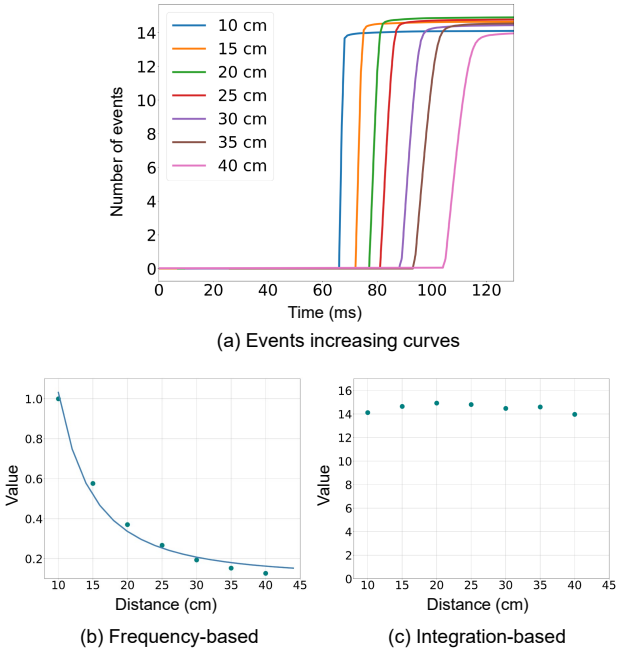

Figure 15. (a) The events increasing curves of settings with lights in different distances. (b) The relationship between transient event frequency and distance. (c) The relationship between integration of events and distance.

results in Fig. 16 and Fig. 17.

8.2. Depth sensing

The transient event frequency can be used for depth sensing, as validated by the inverse square law of light fall-off property described in Eq. (10). In Fig. 15 (a), we measure the number of events and plot events increasing curves of settings with lights at different distances. Based on these curves, we provide graphs showing distances on the horizontal axis and values recovered by our frequency-based and integration-based approaches in Fig. 15 (b) and (c). It is obvious that there are small differences among the number of integrated events of different settings. Thus, the integration-based approach cannot accurately provide the change in radiance with distance, making it difficult to estimate the distance information in a scene. In contrast, the values recovered by the frequency-based approach clearly follow the inverse square law of light shown in the solid line in Fig. 15 (b). This comparison demonstrates the superiority of the frequency-based method on distinguishing the small difference of radiance values.

Depth estimation results of several scenes are shown in Fig. 18. By comparison, we can find that the depth information from an image captured by a conventional RGB camera are easily affected by the exposure time. In the first case of Fig. 18, the near two layers suffer from saturation (over-exposure), which makes them hard to distinguish. In contrast, our frequency-based method is not affected by exposure effect due to the HDR property of event cameras. The integration baseline cannot clearly distinguish layers at different distances.

8.3. Iso-depth contour reconstruction

The comparisons on iso-depth contour reconstruction are shown in Fig. 19. The iso-depth contour maps reconstructed from the integration baseline have lower resolution of radiance values, making it difficult to cluster different points at the object surface into the iso-depth contours. The iso-depth contours recovered by TEF are clear, which is helpful in detecting the shape of the objects.

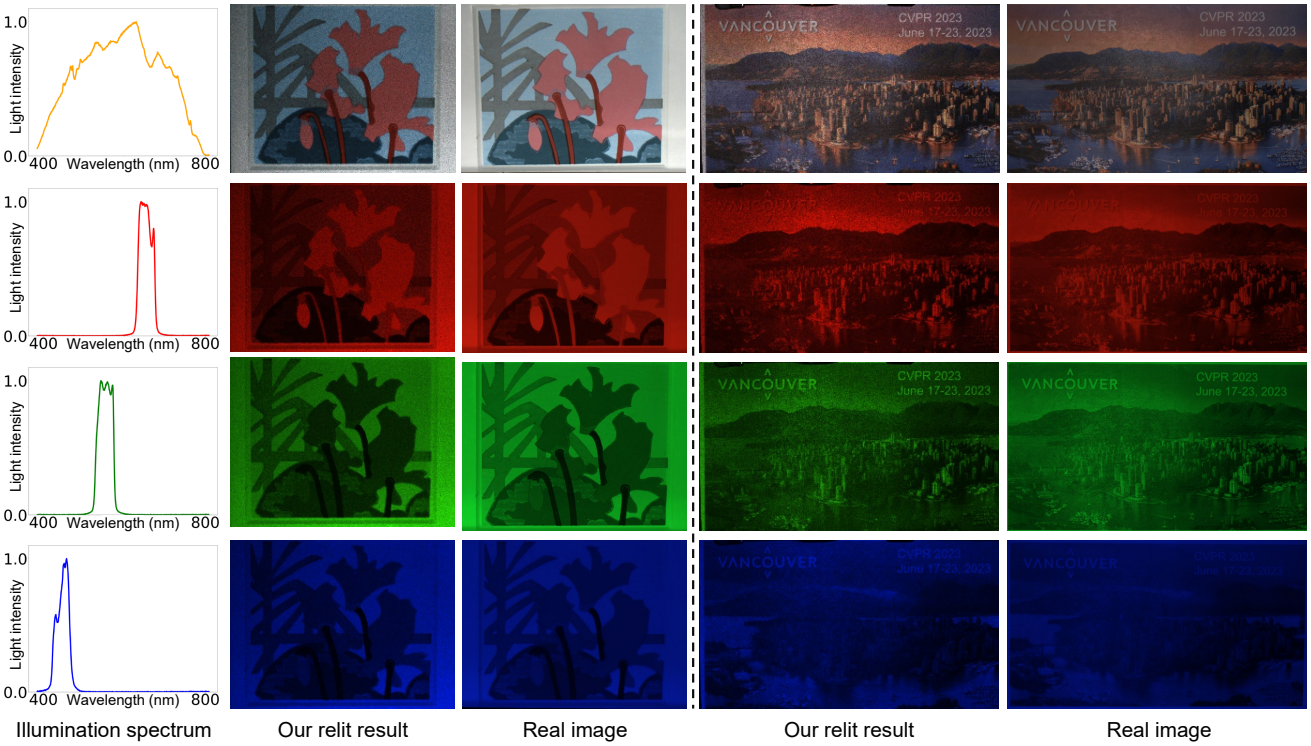

Figure 16. Image relighting results group 1. The relit images are computed from the recovered spectral reflectance curves under 4 different illuminations. The illumination spectrums are measured by a spectrometer. The real images are captured using an RGB camera.

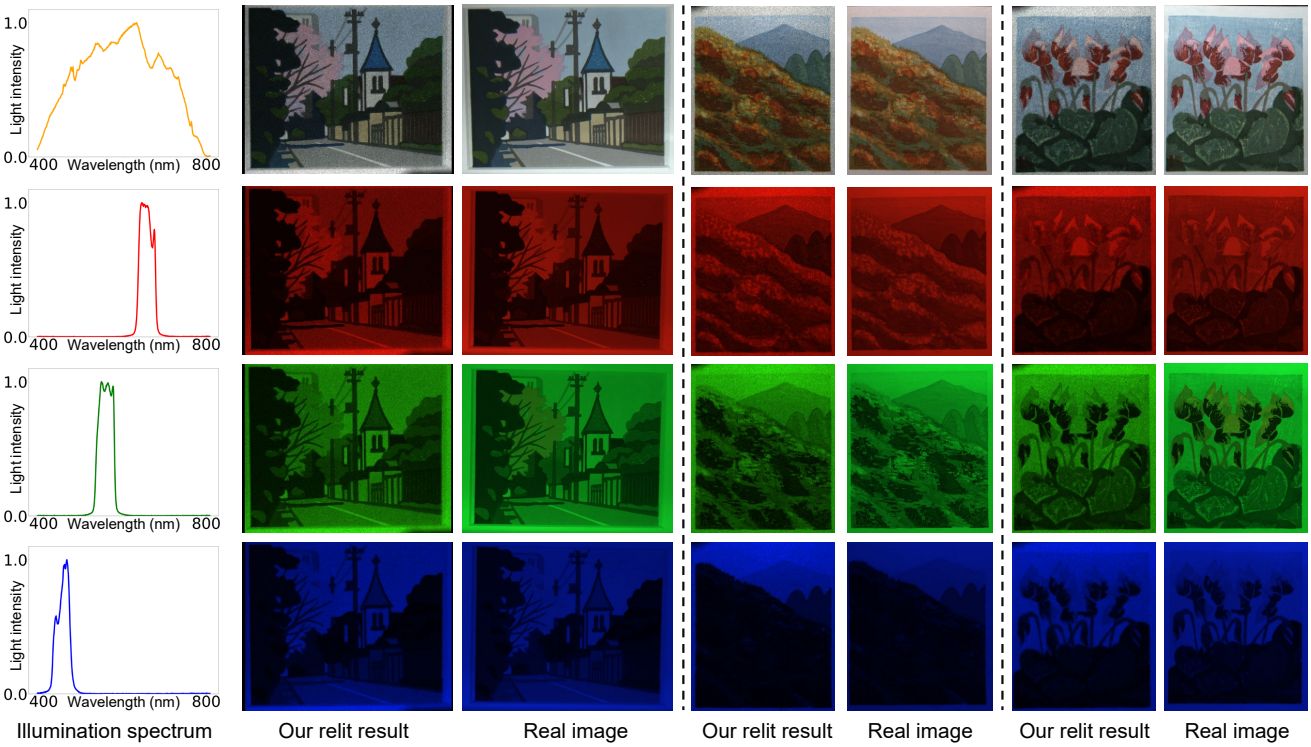

Figure 17. Image relighting results group 2. The relit images are computed from the recovered spectral reflectance curves under 4 different illuminations. The illumination spectrums are measured by a spectrometer. The real images are captured using an RGB camera.

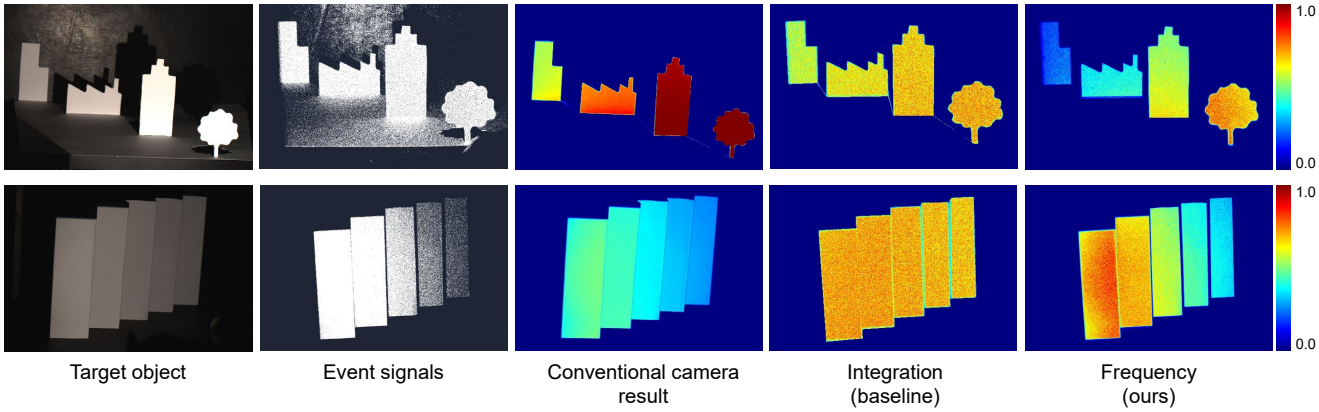

Figure 18. Depth map results comparison. The left 2 columns are raw data captured by a conventional camera and an event camera, respectively. The right 3 columns are depth map results from the conventional camera, the integration of events baseline, and our frequency-based method, respectively.

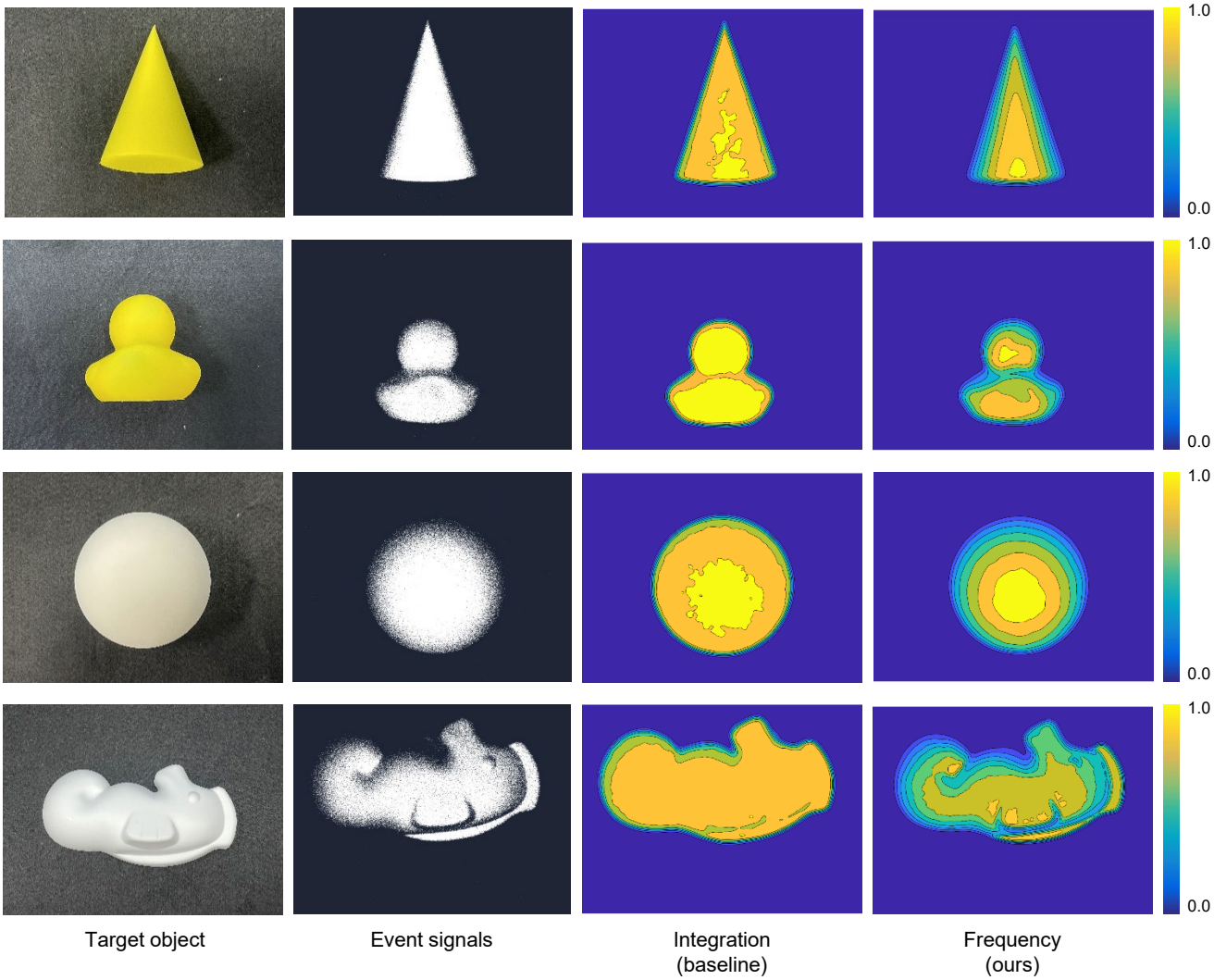

Figure 19. Iso-depth contour results comparison. From left to right columns are target objects captured by a conventional camera, the event signals, and the iso-depth contour reconstruction from integration of events baseline and our frequency-based method.