

A Principle of Event-based Deblurring

According to Eq.1 in paper, the image $I(t)$ can be expressed as Eq.6, where t_r is the reference time. t_r can be any moment within the exposure time $[t_f, t_f + T]$, and $I(t_r)$ is the image at time t_r .

$$I(t_r) \exp(c \int_{t_r}^t e(s) ds) \quad (6)$$

The blurry image is the outcome of integration over the exposure duration $[t_f, t_f + T]$ and is represented as Eq.7.

$$B(t_f) = \frac{1}{T} \int_{t_f}^{t_f+T} I(t) dt \quad (7)$$

According to Eq.6, the blurry image $B(t_f)$ can be transformed into Eq.8.

$$B(t_f) = \frac{I(t_r)}{T} \int_{t_f}^{t_f+T} \exp(c \int_{t_r}^t e(s) ds) dt \quad (8)$$

We define $E(t_r)$ as Eq.9, which is not related to image $I(t)$.

$$E(t_r) = \frac{1}{T} \int_{t_f}^{t_f+T} \exp(c \int_{t_r}^t e(s) ds) dt \quad (9)$$

According to Eq.8 and Eq.9, we can get Eq.10. This demonstrates the correlation among the blurred image $B(t_f)$, the events $e(s)$, and the instantaneous sharp image $I(t_r)$.

$$B(t_f) = E(t_r) \cdot I(t_r) \quad (10)$$

Since actual events are discrete, Eq.9 is transformed into its discrete form, denoted as Eq.11.

$$E[t_r] = \sum_{n=0}^N \exp(c \sum_{i=0}^n e[i]) \quad (11)$$

According to Eq.10 and Eq.11, we can get Eq.(2).

B Details of Dataset

B.1 GOPRO

GOPRO[17] is widely used in deblur-related research and provides ground-truth sharp video. We use the original data to synthesize blurry images and events. To simulate real-world degraded events, we employ v2e[10] for event simulation, as it offers more comprehensive modeling of various characteristics of DVS circuits. This facilitates the simulation of events in diverse environments and under different circuit configurations. During training, blurry images, sharp images, and synthetic events are fed to the model. The training set uses the default parameters in the toolbox to simulate events under undegraded conditions. The circuit simulation parameters used in the test set are entirely different from those in the training set. Specifically, we randomly adjusted several parameters such as threshold variance, shot noise, and cutoff frequency. As shown in the Figure 8, there are two sets of paired images, undegraded events and degraded events with different degrees of degradation. Observably, events characterized by distinct degradation modes exhibit variations in terms of attributes such as noise levels, and edge clarity. Following the suggested training and testing split, the blurry image is also generated by averaging nearby (the number varies from 7 to 13) frames.

B.2 REBlur

The REBlur[29] dataset is designed to provide ground-truth for blurry images through a two-shot approach. Camera motion is controlled by a high-precision motorized slider system, enabling the DAVIS camera to capture pairs of blurry images and sharp images under stable lighting conditions. The dataset comprises 36 sequences and 1469 image pairs, each consisting of two 260×360 grayscale images and events captured during the exposure time of blurry images.

B.3 DavisMCR

The DavisMCR dataset is proposed as part of this study, encompassing a diverse array of degradation events. This dataset comprises 10 lux scenes, with each lux capturing more than 6 objects. In total, the dataset includes 100 sequences, consisting of over 16,000 pairs of images and events. Figure 9 illustrates the data captured under varying lux conditions, ranging from low to high. To simulate different brightness levels in real-world environments, we utilized the ColorSpace CS-HDR-MFS lightbox to create scenes with lux values ranging from 100 to 10000. Figure 10 shows the raw images used to capture DavisMCR data.

To better compare events under different settings, we visualized events captured with a 10ms exposure time. It can be seen from DVS1 in Figure 9 that the ambient brightness has little impact on the event signal-to-noise ratio in the same bright background scene. Comparing DVS1 and DVS2, we can observe that there is more noise in the darker background areas. These events captured in different scenarios are used to evaluate the performance of the deblurring methods.

C Additional Results

C.1 Results of Restoring on DavisMCR Dataset

Consistent with the settings in Section 5.3, we compare the event restoration results of the first stage of RDNet on the DavisMCR dataset with two classic event denoising methods, i.e., SCF[16] and GEF[33]. As shown in Figure 11, (a1) and (c1) are the input images, while (b1) and (d1) are the input original events. (b2-b4) and (d2-d4) show the event restoration results of different methods, and (a2-a4) and (c2-c4) show the deblurred results using the corresponding restored events. To facilitate a more effective comparison of event restoration results, the restored events of SCF and GEF are fed into DeblurNet to obtain deblurred results. The DeblurNet is trained with undegraded events.

As shown in Figure 11, comparing the input events (b1, d1) with the output of SCF (b2, d2), we can see that SCF can only denoise relatively discrete events in space, exhibiting weak event restoration ability. It can also be observed that the reconstruction of GEF relies on the image texture by comparing the input events (b1, d1) with the output of GEF (b3, d3). Therefore, GEF may introduce artifacts by erroneously restoring events with strong textures in non-motion regions. Besides, the comparison between the input events (b1, d1) and the output of RDNet (b4, d4) shows that RDNet can effectively restore events at the motion regions and recover

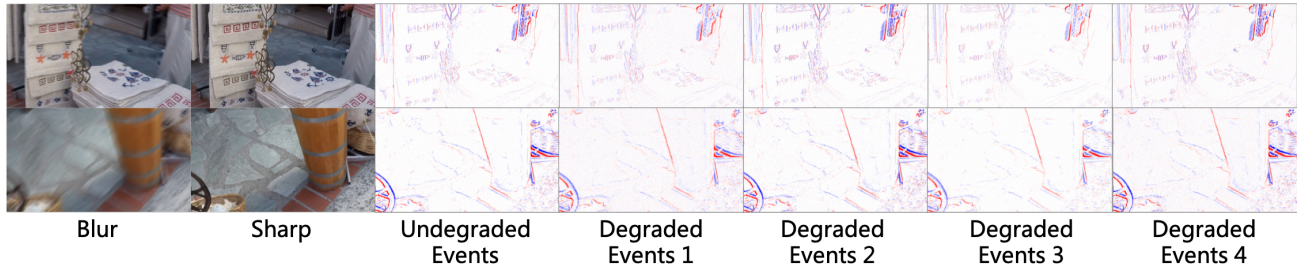


Figure 8: GOPRO training set. Blurry images and sharp images are synthesized from the original dataset. Undegraded events are events simulated with ideal circuit parameters, exhibiting clear motion texture edges and little noise. Degraded events are events simulated with various random degraded circuit parameters, including threshold variance, shot noise, and cutoff frequency.

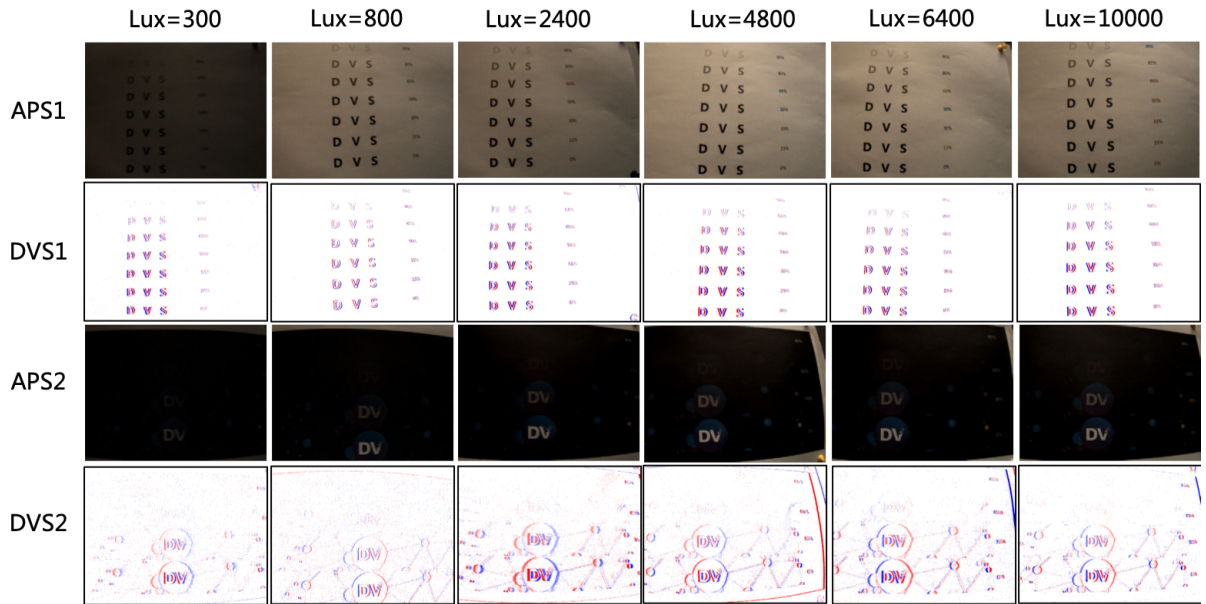


Figure 9: DavisMCR dataset. The columns display images and events captured under different ambient brightness conditions. Distinct ambient brightness levels are typically associated with varying signal-to-noise ratios. APS1 and APS2 represent bright and dark background brightness, respectively. DVS2 captured against a dark background exhibits more noise than DVS1. Objects in different rows within each image have different contrasts. The events in areas with strong contrast are dense and clear.

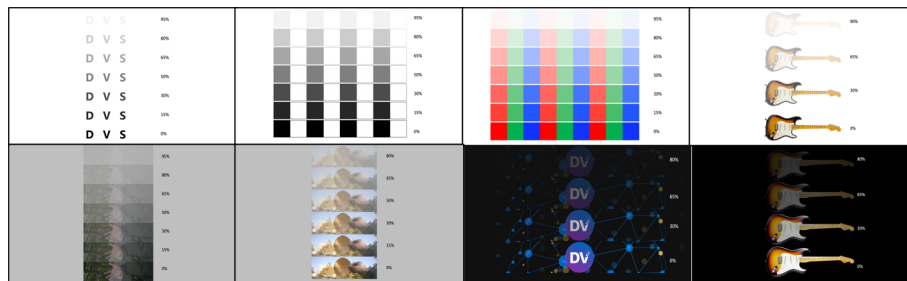


Figure 10: Raw images used to capture DavisMCR data.

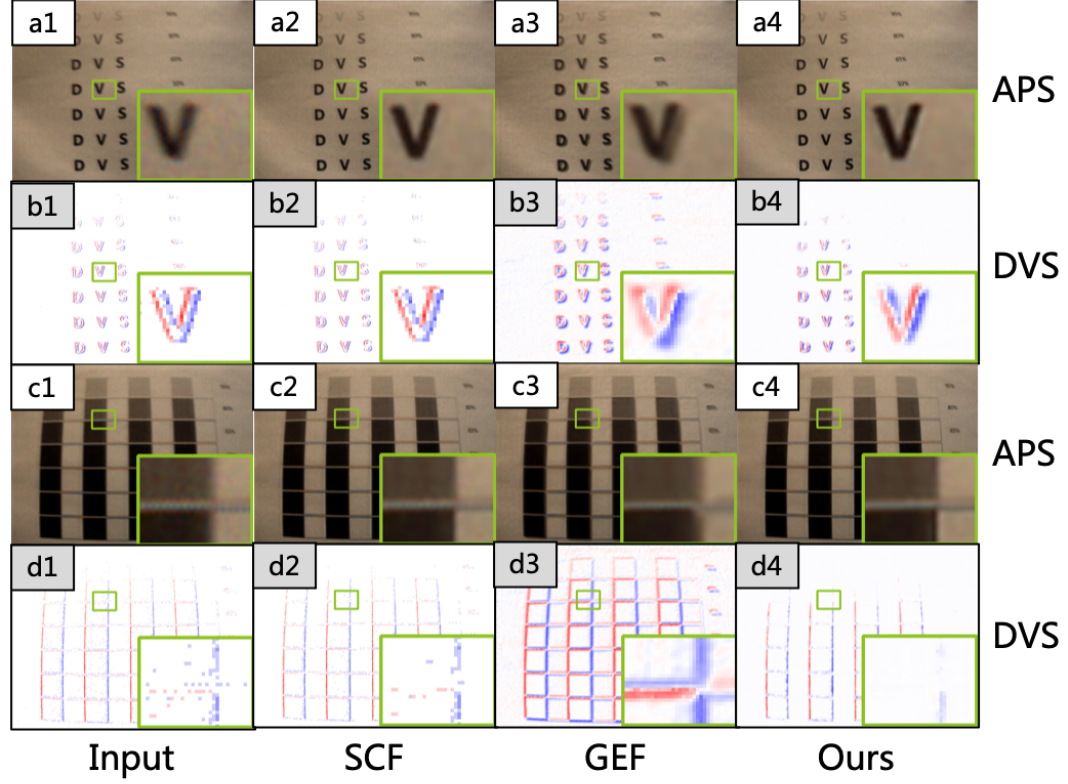


Figure 11: Results of restoring on DavisMCR dataset. (a1) and (c1) represent the input images, while (b1) and (d1) represent the input original events. (b2-b4) and (d2-d4) are the event restoration results of different methods, and (a2-a4) and (c2-c4) are the deblurred results using the corresponding restored events.

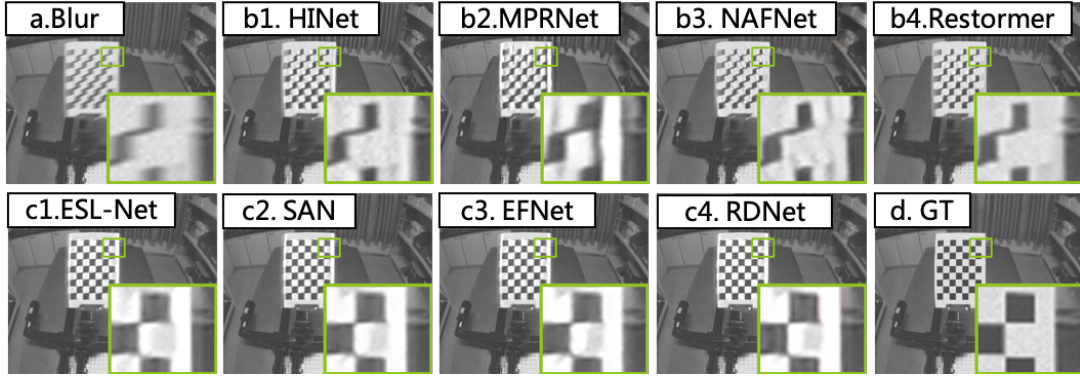


Figure 12: Result of deblurring on REBlur dataset. (a) is the input blurry image, (b*) are the results of image-only deblurring methods, (c*) are the results of event-based deblurring methods, and (d) serves as the ground-truth sharp image.

events in a smooth, high-quality fashion. RDNet, based on the high-quality events restored in the first stage, achieves deblurred results with fewer residual motion blur and artifacts.

C.2 Results of Deblurring on REBlur Dataset

In Figure 12, we compare the qualitative results of deblurring on REBlur. Here, (a) represents the input blurry image, (b*) shows the results of image-only deblurring methods, (c*) demonstrates the

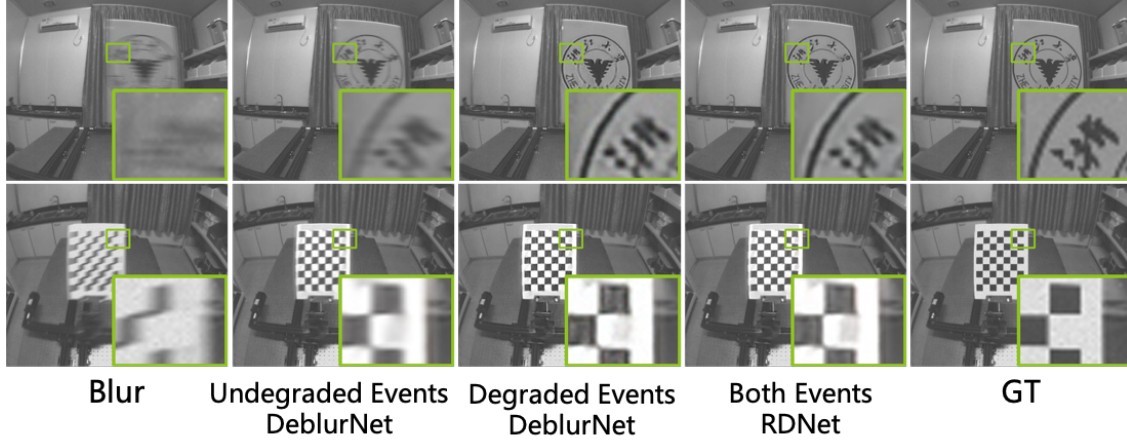


Figure 13: Result of ablation study on REBlur dataset. DeblurNet trained with undegraded events exhibits some residual blur in motion areas, while DeblurNet trained with degraded events generates white-edge artifacts. RDNet trained with both degraded and undegraded events generates the most natural deblurring results, which are closest to the ground-truth.

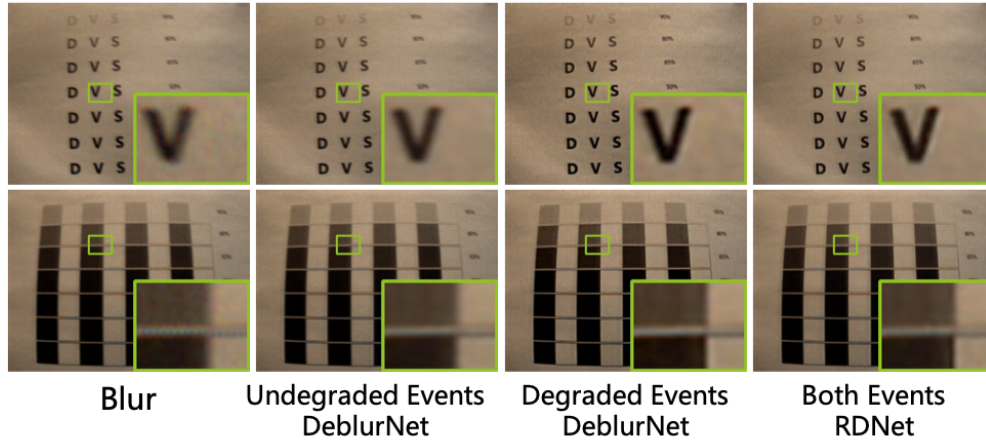


Figure 14: Result of ablation study on DavisMCR dataset. DeblurNet trained with undegraded events exhibits some residual blur in motion areas, while DeblurNet trained with degraded events generates a few white edges. RDNet trained with both degraded and undegraded events generates the best deblurring results.

results of event-based deblurring methods, and (d) serves as the ground-truth sharp image.

In this scenario, the grids on moving objects are severely blurred. This leads to inaccuracies in recovering the texture of blurred images by the methods (b*). Specifically, among these methods, the result of (b2) exhibits erroneous black edges, and the outcome of (b3) is characterized by ghosting.

Compared to the aforementioned image-only methods, the event-based methods (c*) exhibit better capability in restoring the shape of the grid. Among them, the result (c4) of RDNet displays less residual motion blur.

C.3 Results of Ablation Study

Figure 13 and Figure 14 are the results of deblurring on the REBlur dataset and the proposed DavisMCR dataset respectively.

The results of DeblurNet trained with undegraded events show significant residual motion blur in text and grids, while the text and grid become clearer as shown by the results of DeblurNet trained with degraded events. This demonstrates that degraded events effectively simulate the degradation pattern of real-world events. However, the results of DeblurNet trained with degraded events still have some white edge artifacts, which shows that the brightness recovery is inaccurate. In contrast, the results of RDNet trained with both the undegraded and degraded events are clearer and more natural.

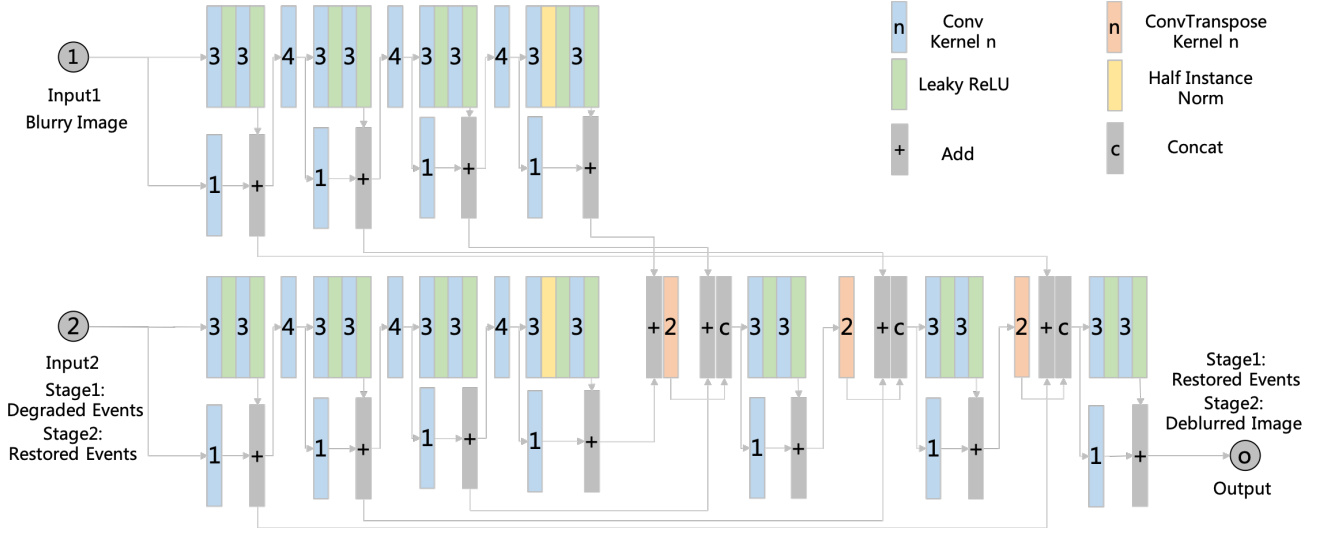


Figure 15: The network structure of each stage of RDNet.

D Structure of RDNet

Figure 15 shows the network structure of each stage of RDNet. This is a network structure with a dual-branch encoder and a single-branch decoder. The network structure of the two stages of RDNet is the same, but their parameters are not shared. Besides, the input

and output of the two stages are different. The input of the first stage is the blurry image and the degraded events, and the output is the restored events. The input of the second stage is the blurry image and the restored events, and the output is the deblurred image.