Exploring the Common Appearance-Boundary Adaptation for Nighttime Optical Flow
Supplementary Material

Hanyu Zhou[1], Yi Chang[1],*, Haoyue Liu[1], Wending Yan[2], Yuxing Duan[1], Zhiwei Shi[1], Luxin Yan[1]
[1]National Key Lab of Multispectral Information Intelligent Processing Technology, School of Artificial Intelligence and Automation, Huazhong University of Science and Technology
[2]Huawei International Co. Ltd.
`{hyzhou,yichang,yanluxin}@hust.edu.cn`

In this supplementary material, we provide the detailed approaches to obtain the spatially aligned nighttime image and event data in Sec. 1. We further provide the details of common space adaptation on improving optical flow during the training process in Sec. 2.1, and the details of common space that plays as the intermediate bridge on domain adaption in Sec. 2.2. Then, we present the generalization of the proposed framework for unseen nighttime scenes in Sec. 3.1 and the robustness of the proposed framework for different illumination conditions in Sec. 3.2, and validate that the proposed method could handle the low light scenes in Sec. 3.3 and high speed scenes in Table 3.4. We also provide the knowledge transfer ability of common space adaptation in Sec. 3.5, and the inference efficiency of the final nighttime optical flow model in Sec. 3.6. We provide the training details of the proposed framework in Sec. 3.7. Besides, we describe the limitation of the proposed nighttime optical flow method in Sec. 3.8. Finally, we provide the qualitative comparison visualization of frame-based optical flow and event-based optical flow models in Sec. 4.

1 Spatial Alignment between Nighttime Image and Event

The prerequisite for boundary adaptation is to obtain the paired nighttime image and event data. The spatially aligned nighttime image and event can be obtained in two ways. First, we have set up a physically coaxial optical system with a beam splitter for the event and image sensor, in which the two modalities are inherently spatial aligned by the shared light path in Fig. 1 (a). Second, we start from the spatial alignment algorithm perspective by performing a standard stereo rectification, and then fine-tune the registration error by pixel offset (Tulyakov et al., 2022) which can ensure a reliable paired nighttime image and event in Fig. 1 (b). It is worth mentioning that, the boundary adaptation is very robust to slight spatial misalignment, since we have introduced the valid mask V generated by the attention map that can distinguish the abnormal regions.

2 Ablation Study

2.1 How does Common Space Adaptation Affect Training Process?

In Fig. 2, we visualize the training process of different space adaptations. Without any domain adaptation in Fig. 2 (a), we train the nighttime optical flow model in an unsupervised manner, where the training curve fluctuates wildly and optical flow result suffers from degradation. This is because nighttime degradation breaks the optical flow basic assumption (Yu et al., 2016), thus nighttime optical flow network cannot effectively learn motion features. In Fig. 2 (b), motion adaptation bypasses the difficulty of directly estimating optical flow from nighttime images, and transfers knowledge from daytime to nighttime domain via domain adaptation in motion space, significantly improving the performance. However, there exist some outliers in the optical flow result due to the large domain gap. In contrast, common appearance adaptation (seeing Fig. 2 (c)) makes the training process more robust and improves results, especially smoothing global motion. Common boundary adaptation (seeing Fig. 2 (d)) further increases the upper limits of optical flow network, and refines the local
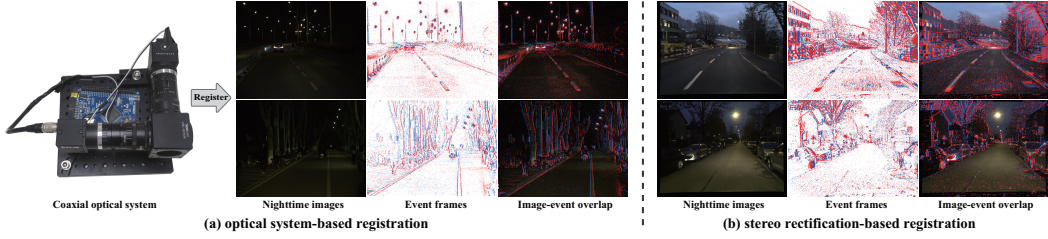
---

*Corresponding author.

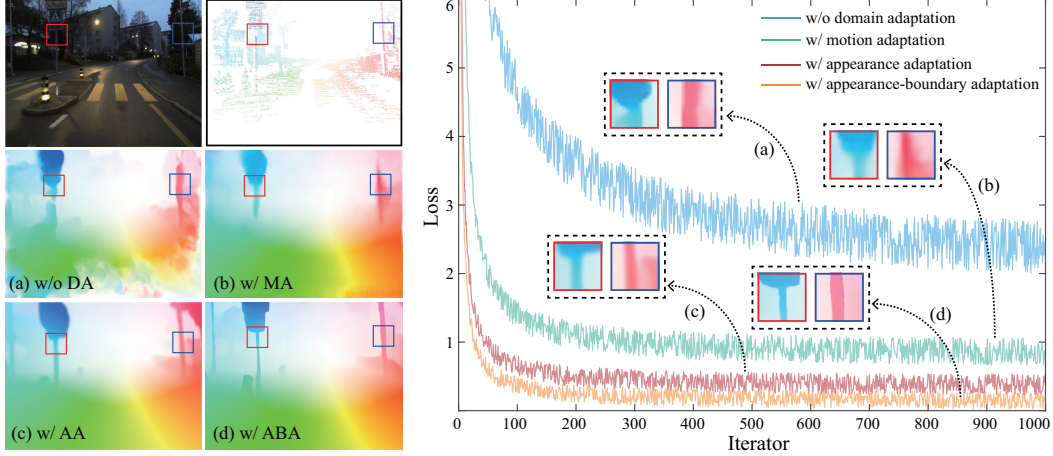Figure 1: Different registration methods for images and events.

Figure 2: Visualization of training process during different space adaptations. Motion adaptation can speed up the convergence of the training curve, and remove most of degradation while there exist obvious outliers. Appearance-boundary adaptation further optimizes the optimal values of the training process, smooths global motion and sharpen local boundary.

motion boundary. Overall, common space adaptation benefits to directionally transfer knowledge to nighttime domain, thus stably training optical flow network and rapidly converging the training curve.

2.2 How does Common Space Serve as the Intermediate Bridge?

In Fig. 3, we illustrate four strategies that common reflectance and boundary spaces play as the intermediate bridges on domain adaptation. In Fig. 3 (a), we directly map visual features to motion space of daytime, nighttime, and event domains, and then transfer motion knowledge to nighttime domain via domain adaptation. However, as shown in Fig. 3 (e), the corresponding optical flow suffers from some outliers due to the large domain gap caused by the intrinsic heterogeneous nature of feature representation (feature distribution misalignment caused by degradation) between daytime/event and nighttime domains. In Fig. 3 (b), we introduce common reflectance space as the intermediate bridge to associate daytime domain and nighttime domain, where we map the visual features of both domains to the common reflectance space and then align their common features. This method could transfer motion appearance knowledge from daytime to nighttime domain, thus smoothing the global motion while suffering from outliers in the boundary regions. Similarly, in Fig. 3 (c), we take common boundary space as an intermediate bridge to associate nighttime domain and event domain, and transfer local boundary knowledge from event to nighttime image domain. The corresponding optical flow is boundary-sharp, but there exist obvious invalid motion regions (seeing the upper-left-corner region of optical flow). Lastly, we choose common reflectance and boundary space as two intermediate bridges in Fig. 3 (d), thus transferring global motion appearance knowledge from daytime to nighttime domain, and local motion boundary knowledge from event to nighttime image domain. The estimated optical flow is global-smooth and boundary-sharp in Fig. 3 (e).

2

(a) Motion adaptation

(b) Appearance-motion adaptation

(c) Motion-boundary adaptation

(d) Appearance-boundary adaptation

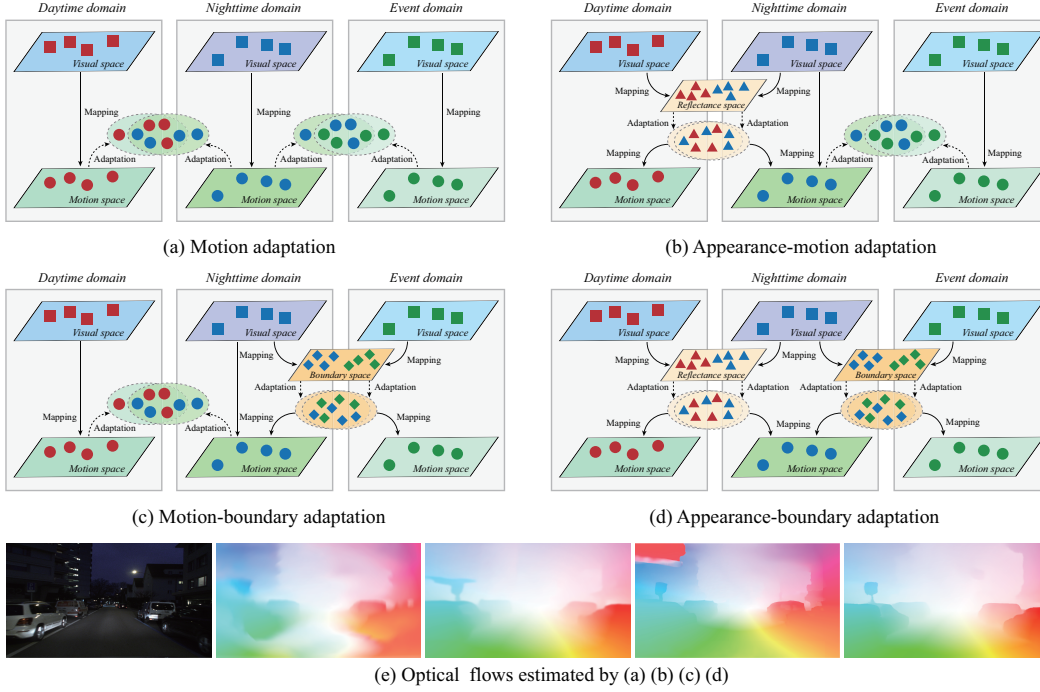(e) Optical flows estimated by (a) (b) (c) (d)

Figure 3: Different strategies of common space playing as the intermediate bridge. There are obvious outliers in the optical flow estimated by (a) motion adaptation. (b) Appearance-motion adaptation with common reflectance space can smooth global motion while the motion boundary suffers from artifacts. (c) Motion-boundary adaptation with common boundary space sharpens motion boundary while there exist patch-level abnormal regions. (d) Appearance-boundary adaptation with common reflectance and boundary spaces can not only smooth global motion, but also sharpen local boundaries.

3 DISCUSSION

3.1 GENERALIZATION FOR UNSEEN NIGHTTIME SCENES

In Fig. 4, in order to compare the generalization of our method with the competing methods for unseen nighttime images, we construct the proposed low light frame-event (LLFE) dataset for generalization experiments, where we have set up a coaxial optical system with a beam splitter for the event and image sensor (seeing Fig. 1 (a)) to simultaneously collect the paired image and event. We can conclude that, the proposed method generalizes well for unseen nighttime scenes with various illumination and consistently outperforms the competing methods.

3.2 ROBUSTNESS FOR VARIOUS ILLUMINATION CONDITIONS

In our framework, common reflectance space is opposite to illumination via retinex model (Fu et al., 2017; Zhu et al., 2017), thus is robust for various illumination conditions. To demonstrate this, we conduct the comparison experiments about optical flow performance of state-of-the-art methods and our method under daytime, dusk and dark conditions in Fig. 5. As for the daytime scenes, all the optical flow methods work well. When applied in dust scenes, the SOTA method GMA (Jiang et al., 2021) suffers from artifacts since degradation from dust scenes weakens visual features, thus matching the inaccurate motion features. The direct visual adaptation method DarkFlow (Zheng et al., 2020) also loses some motion details, while our method still performs well. As for dark conditions, there exist obvious outliers in the optical flows estimated by GMA and DarkFlow. On the contrary, the proposed common space adaptation still works well. Therefore, common space adaptation not only inherits daytime motion knowledge, but also is robust for nighttime scenes with various illumination.
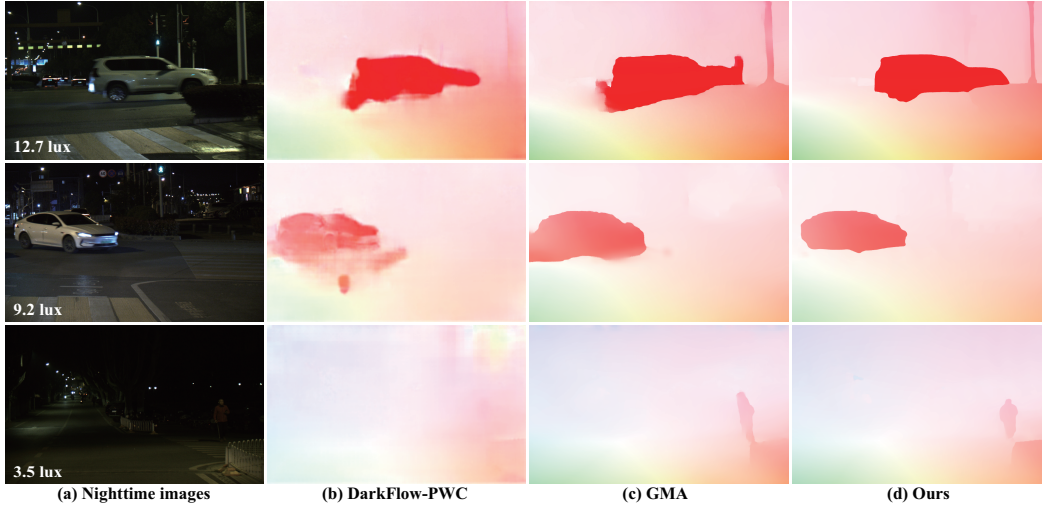
Figure 4: Visualization of comparison on unseen nighttime images with various illuminations.

3.3 SEEING MOTION UNDER EXTREMELY LOW LIGHT SCENES

In this work, we exploit event camera (Gallego et al., 2019) with the advantage of high dynamic range to make up for conventional camera under nighttime scenes, and propose a common boundary adaptation to transfer boundary knowledge from event domain to nighttime image domain. In Fig. 6, we show the effectiveness of common boundary adaptation for extremely low light conditions. The state-of-the-art optical method GMA and the direct visual adaptation method DarkFlow cannot almost estimate the optical flow, while the proposed common boundary adaptation still performs well under extremely low light conditions. The main reason is, event camera could see motion boundaries that conventional camera cannot see in extremely low-light regions. Hence, the common boundary adaptation transfers local boundary knowledge from event domain to nighttime image domain.

3.4 APPLICATION FOR LOW LIGHT AND HIGH SPEED SCENARIOS

Low light and high speed are two of the more challenging scenarios for nighttime scenes. In terms of imaging mechanism, conventional frame camera records the absolute luminance with a fixed exposure time via global scan. In nighttime scenes, conventional frame camera would inevitably face a dilemma between the long exposure time of low-light scenarios and motion blur of high-speed scenarios. In contrast, the event camera reacts to changes in light intensity, rather than integrating photons during the exposure time of each frame (Cabriel et al., 2023). Each pixel works independently and returns a signal only when an intensity change is detected. Compared with the conventional frame camera, the event camera can sense the dynamic changes with higher temporal resolution (microsecond) and higher dynamic range, thus compensating for frame camera in nighttime scenes, such as, nighttime image enhancement (Liang et al., 2023) and nighttime deblurring (Qi et al., 2023).

Similarly, the event camera can also assist the frame camera to learn the motion patterns in nighttime low-light and high-speed scenarios. To discuss the impact of the event on the optical flow in nighttime low-light and high-speed scenes, as shown in Table 1, we use the coaxial optical system in Fig. 1 to collect the spatiotemporally-aligned image sequences and events stream with various illumination (*e.g.*, 3.5 lux, 9.2 lux and 12.7 lux) and various driving speed (*e.g.*, 50 km/h, 70 km/h and 80 km/h) to quantitatively compare the optical flow performance. As for the optical flow label, since it is difficult to directly obtain the dense optical flow labels, we manually mark 100 pairs of corresponding corner points for each two adjacent images, and calculate the relative displacement between the corner points as the sparse optical flow labels. In addition, we choose EPE as the evaluation metric. We have two observations. First, the event camera can greatly improve optical flow performance in both nighttime low-light and nighttime high-speed scenes. In high-speed scenes, the proposed method is robust to various speeds, and the optical flow performance remains unchanged, demonstrating the advantage of high temporal resolution of the event camera. In low-light scenes, as the illumination becomes lower, the optical flow metric (EPE) trend becomes larger obviously. This shows that, although the event camera has the advantage of high dynamic range, too low illumination would also interfere with the
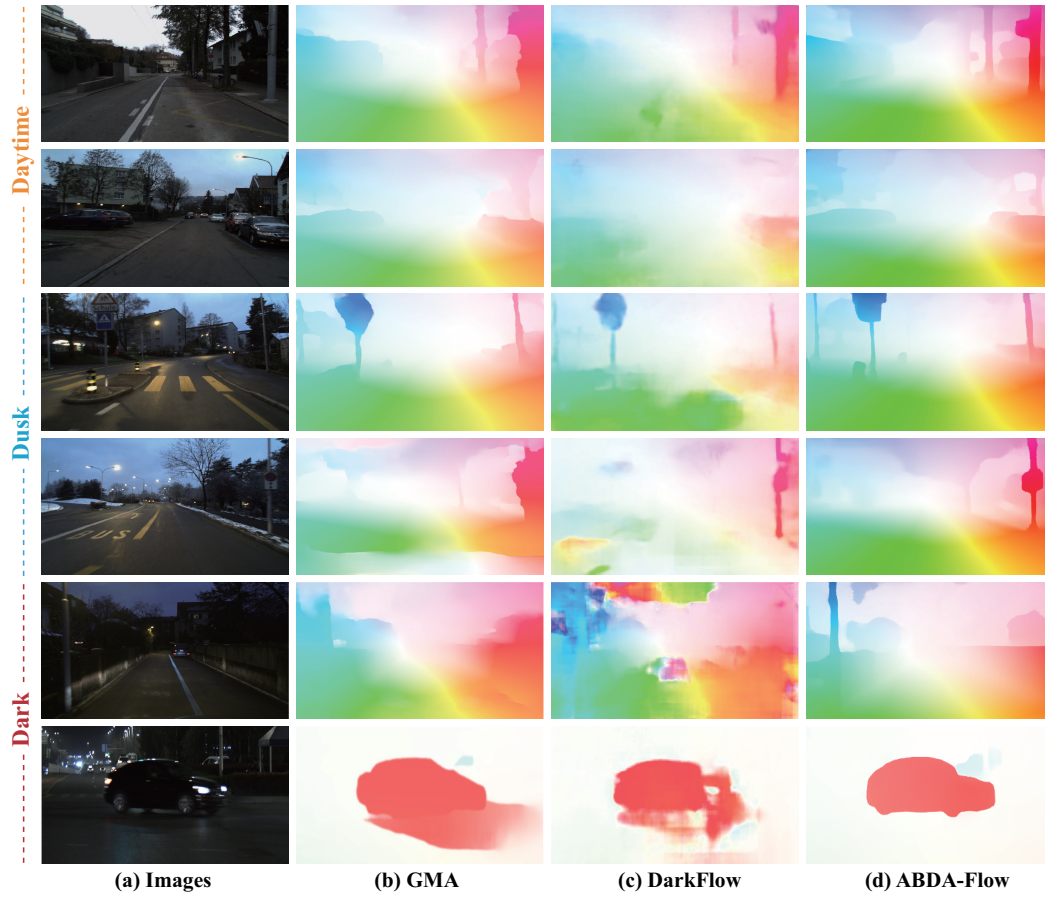
Figure 5: Visualization of optical flow under various illumination conditions.

(a) Images   (b) GMA   (c) DarkFlow   (d) ABDA-Flow

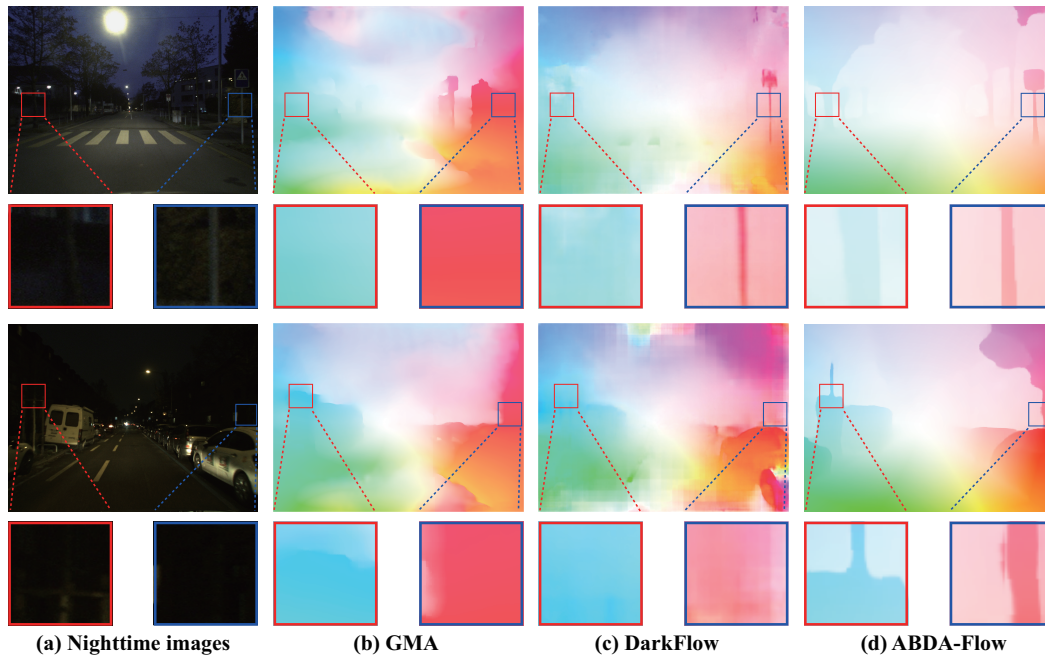(a) Nighttime images   (b) GMA   (c) DarkFlow   (d) ABDA-Flow

Figure 6: Visualization of optical flows under extremely low light conditions.

Table 1: Quantitative results of the proposed method in low light and high speed scenarios.

Method	Low light			High speed		
	3.5 lux	9.2 lux	12.7 lux	50 km/h	70 km/h	80 km/h
Our flow baseline	4.26	3.98	3.67	3.65	4.53	5.74
Our flow baseline, w/ only event	2.05	2.01	1.63	1.75	1.75	1.80
Our flow baseline, w/ only event, w/ common space	1.52	1.44	1.10	1.14	1.15	1.19

Table 2: Discussion on transfer ability of the proposed method from daytime to nighttime domain.

Test Dataset	Daytime images of DSEC	Nighttime images of DSEC
w/o adaptation	10.71%	25.42%
w/ direct motion adaptation	10.71%	18.82%
w/ common reflectance space adaptation	10.71%	14.51%

optical flow performance. The main reason is that, under low light conditions, the event noise is intensified. The event noise and the valid signal event are both 0-1 pulses, and their difference is very small, which affects the optical flow. Second, the proposed common space can further improve the upper limit of optical flow, indicating that common space can serve as a bridge to reinforce the feature alignment between event and nighttime image domains.

3.5 ABILITY OF DAYTIME→NIGHTTIME MOTION KNOWLEDGE TRANSFER

In the proposed framework, the performance of the daytime optical flow is the upper limit of the nighttime optical flow. Thus, the key to estimating nighttime optical flow lies in the ability of motion knowledge transfer from daytime domain to nighttime domain. As shown in Table 2, we conduct an ablation study to analyze the knowledge transfer ability of the common reflectance adaptation. Three strategies use the same pre-trained daytime optical flow model to estimate daytime optical flow, and they have the same performance. When tested on the nighttime images, w/o adaptation performs badly, and w/ direct motion adaptation (namely only motion distribution alignment) performs relatively well, while w/ common reflectance space adaptation (namely appearance adaptation) performs better. In conclusion, the proposed common reflectance space adaptation has a better ability to transfer motion knowledge from daytime domain to nighttime domain.

3.6 INFERENCE TIME

In Table 3, we choose inference time as the efficiency metric for optical flow estimation, and RTX 3090 as the inference platform. We can observe that, RAFT and GMA can satisfy the requirement for real-time application, while their performances are bad under nighttime conditions. In contrast, the proposed method can estimate accurate optical flow, while the inference time is not efficient enough due to the complex transformer as the optical flow backbone. Moreover, under the common adaptation framework, we replace the transformer with GMA as the optical flow backbone, namely GMA w/ common adaptation, which performs well and efficiently on nighttime images. This shows that our common adaptation framework has the advantages of universality and easy scalability without increasing processing time. Therefore, our method can further choose an efficient and expressive backbone to improve processing efficiency.

3.7 IMPLEMENTATION OF TRAINING DETAILS

Contrastive Learning Sample Number Setup. In Table 4, we study the influence of the contrastive sample number on the final optical flow result. We select $[100, 500, 1000, 2000]$ as the candidate values of sample number N. We can observe that the larger the sample number N, the better the optical flow result is. However, when the sample number N is increased to 2000, optical flow metric is not improved much, but instead, the computation cost is increased. Hence, to make a trade-off between performance and cost, we set 1000 as the feature sample number N of contrastive learning.

Table 3: Discussion on inference time.

Method	Inference time (ms)	EPE	Fl-all
RAFT	54	1.69	28.65%
GMA	79	1.48	24.10%
GMA w/ common adaptation	79	0.79	12.39%
Transformer w/ common adaptation (ours)	273	**0.74**	**11.85%**

Table 4: Discussion on contrastive sample number.

Sample number	EPE	Fl-all
100	0.81	13.60%
500	0.77	12.26%
1000	**0.74**	11.85%
2000	**0.74**	**11.78%**

Table 5: The choice of motion class number.

Motion classes	EPE	Fl-all
2	0.81	13.42%
5	0.76	12.10%
10	**0.74**	**11.85%**
15	0.79	12.54%

How to Choose Motion Class Number? Motion class number K is a parameter that measures the degree of motion degradation. As shown in Table 5, the motion class number is not as more as possible, but there is a balance, namely 10. The reason is that motion classification depends on the probability threshold, the larger the motion class number value, the larger the probability threshold corresponding to the normal motion feature, increasing the risk of misclassification of abnormal motion features. Therefore, an appropriate motion class number is important to the final result.

Weight Sensitivity of Model Losses. To choose the optimal weight parameters for the total loss, we conduct the study on the weight sensitivity of the typical adaptation losses in Fig. 7, such as, $\mathcal{L}_{intra}^{align}$, $\mathcal{L}_{inter}^{align}$, $\mathcal{L}_{contra}$ and $\mathcal{L}_{flow}^{self}$. $\mathcal{L}_{intra}^{align}$ aims to transfer knowledge between visual-based motion space and reflectance-based motion space within the same domain. $\mathcal{L}_{inter}^{align}$ is to transfer motion knowledge between daytime and nighttime domains. The intention of $\mathcal{L}_{contra}$ and $\mathcal{L}_{flow}^{self}$ is to transfer motion knowledge from event to nighttime image domain along the motion feature dimension and motion field dimension, respectively. In Fig. 7 (a), the larger the weight of $\mathcal{L}_{intra}^{align}$, the more rapidly the optical flow network converges. In Fig. 7 (b), the K-L divergence loss $\mathcal{L}_{intra}^{align}$ is sensitive to the framework training. If the weight is too large, the gradient will disappear. In Fig. 7 (c), a larger weight of the loss $\mathcal{L}_{contra}$ instead contributes negatively to the final optical flow results. The main reason is that a too large weight may make the network pay too much attention to contrastive learning and ignore other discriminative features. In Fig. 7 (d), the weight of $\mathcal{L}_{flow}^{self}$ is robust for the framework training. Therefore, we set the adaptation losses weights as $[\lambda_3, \lambda_4, \lambda_6, \lambda_7]$ as $[1.0, 0.01, 0.01, 1.0]$.

3.8 LIMITATION

The main limitation is that the proposed method cannot accurately estimate the radial motion. There are two reasons. First, the optical flow network usually estimates the motion along the x, y-axis direction of the camera plane, while there is no relative x, y-axis motion when the radial moving object is projected onto the camera plane. Second, for the event camera, the radial moving objects only trigger a small number of events, which affects the event optical flow estimation. For example, when a vehicle is approaching us from far to near along the center of the camera, there is no relative motion in the frame camera plane and no triggered events corresponding to this vehicle in the event camera. In the future, we will integrate lidar into the proposed framework to estimate radial motion.

4 COMPARISON

4.1 COMPARISON ON SYNTHETIC IMAGES

The visual results of optical flow predicted by ABDA-flow and other state-of-the-art approaches on the synthetic Noise Dark-KITTI2015 dataset are presented in Figure 8. The competing methods include SMURF (Stone et al., 2021) and DarkFlow (Zheng et al., 2020). For the SMURF method, we
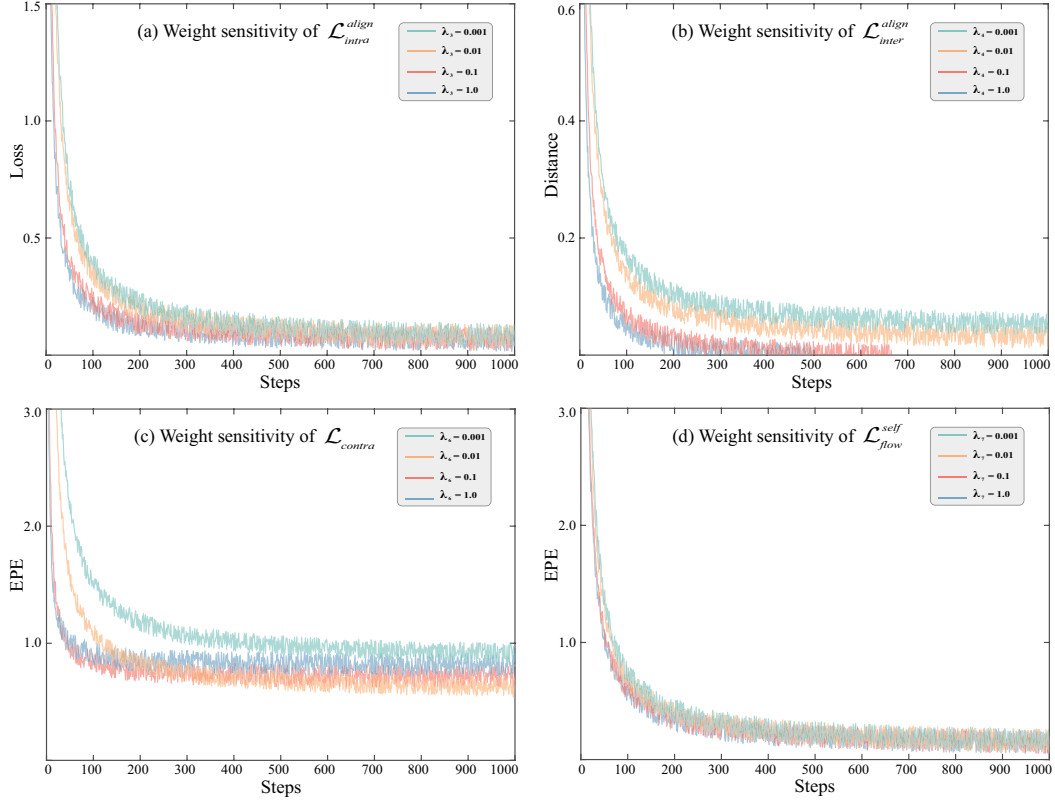
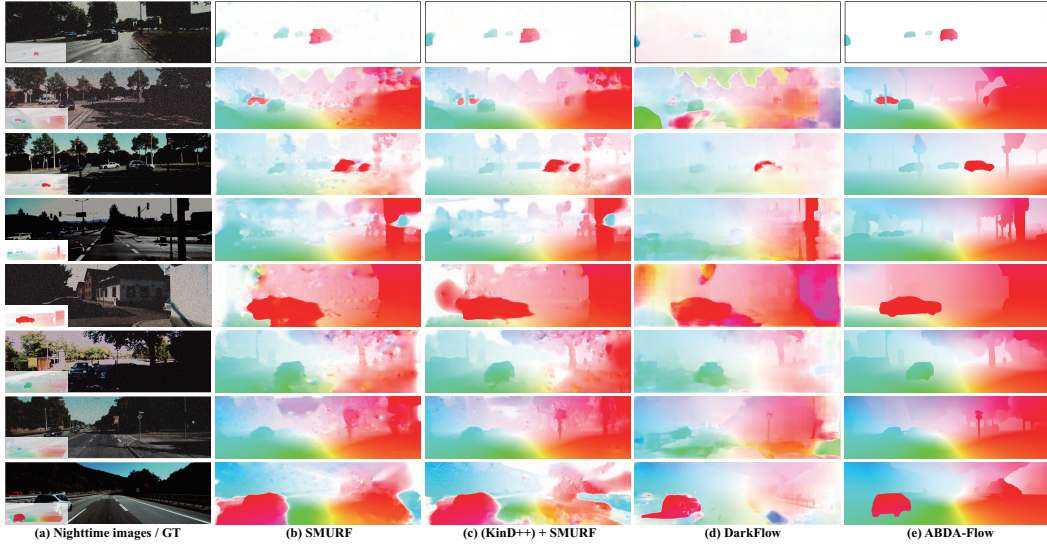Figure 7: The weight sensitivity of model adaptation losses.

Figure 8: Comparison of optical flows on synthetic Noise Dark-KITTI2015 dataset.

utilize the KinD++ (Zhang et al., 2019) image enhancement technique, applying it to the enhanced results of Noise Dark-KITTI2015 and denoting it as (KinD++) + SMURF. Upon observation, it is evident that the optical flows estimated by SMURF exhibit noticeable artifacts, primarily due to the violation of the optical flow basic assumption in nighttime scenes. Although (KinD++)+SMURF reduces the degradation to some extent, it does not function properly. This can be attributed to the fact that the image enhancement method KinD++ is not specifically designed for nighttime optical flow. As a result, any residual artifacts may smooth out the visual features, thereby compromising the accuracy of motion feature matching. While DarkFlow significantly improves nighttime optical flow,
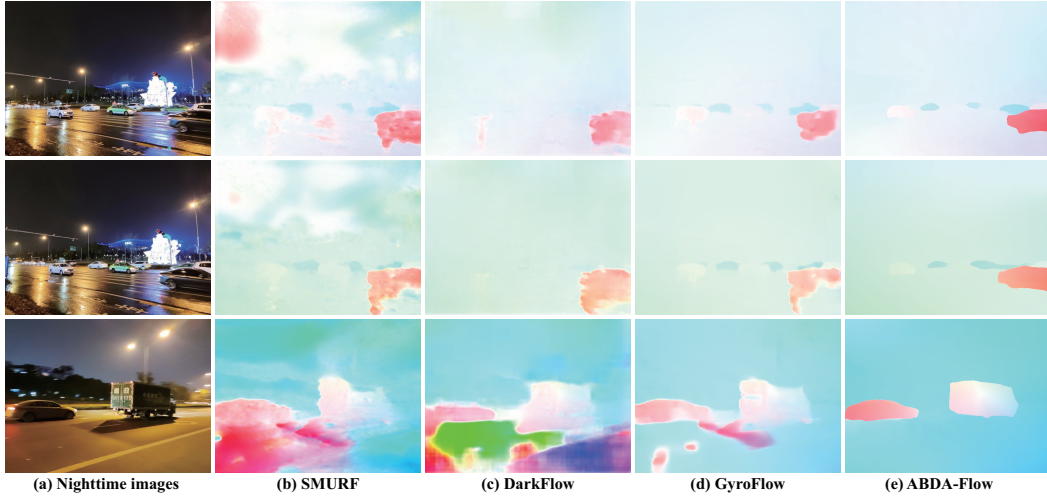
8

| (a) Nighttime images | (b) SMURF | (c) DarkFlow | (d) GyroFlow | (e) ABDA-Flow |

Figure 9: Comparison of optical flows on real Dark-GOF dataset.

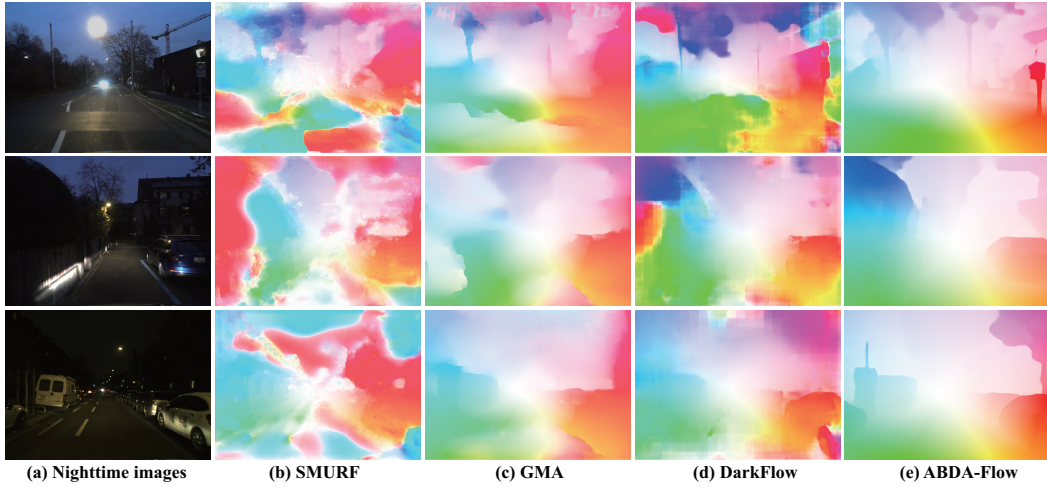| (a) Nighttime images | (b) SMURF | (c) GMA | (d) DarkFlow | (e) ABDA-Flow |

Figure 10: Comparison of optical flows on real Dark-DSEC dataset.

slight outliers still exist due to the substantial domain gap between daytime and nighttime domains. In contrast, the common space adaptation approach mitigates the misalignment between both domains, enabling the estimation of accurate optical flow.

4.2 COMPARISON ON REAL IMAGES

Comparison on Dark-GOF dataset. In Fig. 9, we also provide the visualization of optical flows estimated by the unsupervised method SMURF (Stone et al., 2021), visual adaptation DarkFlow (Zheng et al., 2020), motion adaptation GyroFlow (Li et al., 2021) and our common space adaptation ABDA-Flow on the Dark-GOF dataset. SMURF cannot work normally since nighttime degradation breaks the optical flow basic assumption. Visual adaptation DarkFlow can remove some artifacts, but there exist obvious outliers in the optical flow results. Motion adaptation GyroFlow smooths the global motion, while failing for independent motion objects. This direct adaptation in visual space and motion space is ineffective due to the large domain gap since there exists the intrinsic heterogeneous nature (feature distribution misalignment caused by degradation) of feature representations. However, the proposed ABDA-Flow performs well under real nighttime scenes since it takes the common-latent space to reinforce the feature alignment.

Comparison on Dark-DSEC dataset. We also show the results of our proposed ABDA-Flow on real Dark-DSEC images in Fig. 10, in which there are more complex motion patterns and nighttime conditions. The unsupervised method SMURF hardly works since nighttime degradation breaks the

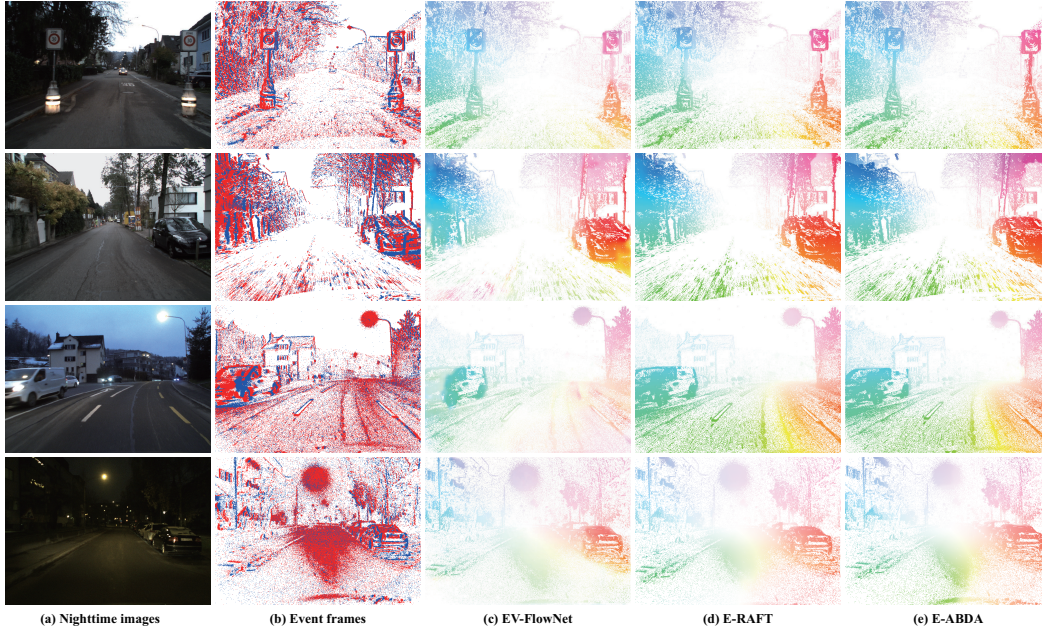| (a) Nighttime images | (b) Event frames | (c) EV-FlowNet | (d) E-RAFT | (e) E-ABDA |

Figure 11: Comparison of event-based optical flows on real event stream from DSEC dataset.

optical flow basic assumption which unsupervised methods rely on. The supervised method GMA cannot work well due to the domain gap between synthetic and real images. There are some artifacts in the optical flow estimated by the visual adaptation DarkFlow. The main reason is that there exists a large gap between daytime and nighttime domains, in which the feature misalignment in visual space causes inaccurate feature matching in motion space. Note that, since these three methods estimate optical flow via frames, their performances are limited due to the low dynamic range of conventional camera under extremely low light conditions (seeing the fourth row in Fig. 10). On the contrary, ABDA-Flow still works better on complex nighttime images, even if extremely low light.

4.3 COMPARISON ON EVENT OPTICAL FLOW

In Fig. 11, we compare the state-of-the-art event optical flow models (EV-FlowNet (Zhu et al., 2018) and E-RAFT (Gehrig et al., 2021b)) with our event model on the real event stream from DSEC (Gehrig et al., 2021a) dataset. We can observe that the optical flow estimated by EV-FlowNet is over-smooth, and E-RAFT losses slight motion details in the motion boundaries. Instead, our event optical flow E-ABDA still works well, verifying its superiority.

REFERENCES

Clément Cabriel, Tual Monfort, Christian G Specht, and Ignacio Izeddin. Event-based vision sensor for fast and dense single-molecule localization microscopy. *Nature Photonics*, 17(12):1105–1113, 2023. 4

Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John Paisley. Removing rain from single images via a deep detail network. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 3855–3863, 2017. 3

Guillermo Gallego, Tobi Delbrück, G. Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew J. Davison, Jörg Conradt, Kostas Daniilidis, and Davide Scaramuzza. Event-based vision: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, 44:154–180, 2019. 4

Mathias Gehrig, Willem Aarents, Daniel Gehrig, and Davide Scaramuzza. DSEC: A stereo event camera dataset for driving scenarios. *IEEE Robotics and Automation Letters*, 2021a. 10

Mathias Gehrig, Mario Millhäusler, Daniel Gehrig, and Davide Scaramuzza. E-raft: Dense optical flow from event cameras. In *International Conference on 3D Vision*, pp. 197–206, 2021b. 10

Shihao Jiang, Dylan Campbell, Yao Lu, Hongdong Li, and Richard Hartley. Learning to estimate hidden motions with global motion aggregation. In *Int. Conf. Comput. Vis.*, pp. 9772–9781, 2021. 3

Haipeng Li, Kunming Luo, and Shuaicheng Liu. Gyroflow: Gyroscope-guided unsupervised optical flow learning. pp. 12869–12878, 2021. 9

Jinxiu Liang, Yixin Yang, Boyu Li, Peiqi Duan, Yong Xu, and Boxin Shi. Coherent event guided low-light video enhancement. In *Int. Conf. Comput. Vis.*, pp. 10615–10625, 2023. 4

Yunshan Qi, Lin Zhu, Yu Zhang, and Jia Li. E2nerf: Event enhanced neural radiance fields from blurry images. In *Int. Conf. Comput. Vis.*, pp. 13254–13264, 2023. 4

Austin Stone, Daniel Maurer, Alper Ayvaci, Anelia Angelova, and Rico Jonschkowski. Smurf: Self-teaching multi-frame unsupervised raft with full-image warping. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 3887–3896, 2021. 7, 9

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 *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 17755–17764, 2022. 1

Jason J Yu, Adam W Harley, and Konstantinos G Derpanis. Back to basics: Unsupervised learning of optical flow via brightness constancy and motion smoothness. In *Eur. Conf. Comput. Vis.*, pp. 3–10, 2016. 1

Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer. In *ACM Int. Conf. Multimedia*, pp. 1632–1640, 2019. 8

Yinqiang Zheng, Mingfang Zhang, and Feng Lu. Optical flow in the dark. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 6749–6757, 2020. 3, 7, 9

Alex Zihao Zhu, Liangzhe Yuan, Kenneth Chaney, and Kostas Daniilidis. Ev-flownet: Self-supervised optical flow estimation for event-based cameras. *Robotics: Science and Systems*, 2018. 10

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Int. Conf. Comput. Vis.*, pp. 2223–2232, 2017. 3