Optical Flow in the Dark

Yinqiang Zheng[1,*], Mingfang Zhang[2,3,*], Feng Lu[2,4,†]
[1]National Institute of Informatics, [2]Pengcheng Laboratory, [3]Nanjing University,
[4]State Key Lab. of VR Technology and Systems, School of CSE, Beihang University
yqzheng@nii.ac.jp, 161220169@smail.nju.edu.cn, lufeng@buaa.edu.cn

Abstract

Many successful optical flow estimation methods have been proposed, but they become invalid when tested in dark scenes because low-light scenarios are not considered when they are designed and current optical flow benchmark datasets lack low-light samples. Even if we preprocess to enhance the dark images, which achieves great visual perception, it still leads to poor optical flow results or even worse ones, because information like motion consistency may be broken while enhancing. We propose an end-to-end data-driven method that avoids error accumulation and learns optical flow directly from low-light noisy images. Specifically, we develop a method to synthesize large-scale low-light optical flow datasets by simulating the noise model on dark raw images. We also collect a new optical flow dataset in raw format with a large range of exposure to be used as a benchmark. The models trained on our synthetic dataset can relatively maintain optical flow accuracy as the image brightness descends and they outperform the existing methods greatly on low-light images.

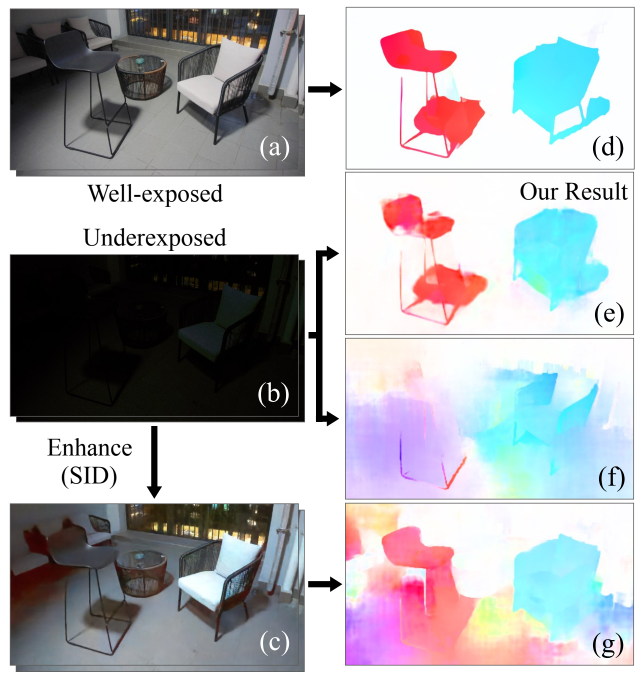


Figure 1. A comparison between our method (e) and existing methods (d,f,g), conducted on the well-exposed input (a), underexposed input (b) and the enhanced input (c) using SID [6] method. Specifically, (a) and (b) are images of the same scene. (e,f,g) are produced by PWC-Net [28] and (d) is produced by FlowNet2 [14] which is 17 times larger than PWC-Net.

1. Introduction

Optical flow estimation can be used for motion detection, object segmentation, and so on. In real life, many of the applications need to deal with low-light data. For example, there are certain circumstances for a drone to fly at night but it's difficult to achieve the goal because optical flow is hard to obtain in dark environments, which is useful to avoid obstacles and control the speed through dense surroundings. Current optical flow methods show poor performance with low-light data. To address this problem, a straightforward solution is to use auxiliary lighting systems for drones, which would inevitably shorten the battery life. As a result, an efficient optical flow method to process low-light data is greatly demanded, which is the main goal of our work.

Horn and Schunck [13] introduce an energy minimization approach to compute optical flow and many excellent methods adapt it and achieve better and better results. However, the optimizing problem of a complex energy function is usually computationally expensive for real-time applications. In order to create a fast and accurate method, end-to-end convolutional neural network frameworks are pro-

*Equal contribution.
†Corresponding author.
This work is partially supported by the National Natural Science Foundation of China (NSFC) under Grant 61972012 and Grant 61732016.
This work is finished during Y. Zheng's visit and M. Zhang's internship at Pengcheng Laboratory.

posed, such as FlowNetS, FlowNetC [10], FlowNet2 [14], and PWC-Net [28]. The best of them can achieve state-of-the-art performance with both better quality and faster speed than traditional methods. As a result, we choose to develop our approach based on deep-learning methods. However, they still show poor capabilities of dealing with low-light data because low-light is a complex scenario due to the low signal-to-noise ratio, and current optical flow datasets, eg. KITTI [12], Sintel [4], FlyingChairs [10] and Flyingthings3D [22], mainly consist of bright images. Low-light is not taken into consideration when current optical flow methods are designed.

A direct solution to get optical flow from dark images is to enhance them before computing optical flow. Researchers have proposed many techniques for denoising and enhancing low-light images [9, 3, 6, 20]. However, the purpose of these techniques is to get better visual quality of the enhanced images, and we get limited improvement on optical flow results or even worse ones. The main reason is that many of the methods lose information while enhancing, such as brightness constancy which is fundamental for optical flow estimation. As a result, we choose to make the network learn optical flow directly from low-light images in an end-to-end way that avoids information loss.

It is known that optical flow ground truth is hard to obtain and neural network training needs a large scale of data, so we decide to synthesize low-light effects on bright images and create low-light optical flow datasets. A variety of learning-based denoising and low-light enhancing methods synthesize their own dataset. Some of them [18, 19] use the additive white noise (AWGN) model, which is always criticized that they are not capable of simulating real noise and recent denoising works [15, 2, 23] propose to simulate camera processing features which leads to better noise models. We focus on simulating the noise on raw images, which avoids complex characteristics due to the image processing pipeline. Most existing noise analysis [2] is conducted on bright noisy images [24] in which the noise is caused by high ISO. Beside that, we mainly focus on the noise caused by the low intensity of light, since the photons arriving at the sensor are governed by Poisson distribution. And finally, we apply our noise model to create new optical flow datasets for training. In order to test a model's ability to deal with various exposure inputs, we collect a new optical flow dataset in low-light environments.

Contribution

1. We collect a new optical flow dataset — Various Brightness Optical Flow (VBOF) dataset, consisting of 598 raw images of various brightness and 100 optical flow references to evaluate the ability of an optical flow model to deal with various exposure inputs.
2. We provide an analysis of the noise on raw images of various brightness and reveal how the noise distribution changes as the exposure descends.
3. We apply our noise model to synthesize a low-light optical flow dataset — FlyingChairs-DarkNoise (FCDN), and the optical flow models trained on our dataset can relatively maintain optical flow accuracy as the exposure descends and they outperform the existing methods greatly on low-light images.

2. Related Work

Learning-based Optical Flow Estimation. Dosovitskiy et al. [10] propose two CNN models for optical flow, FlowNetS and FlowNetC, which shows the feasibility of directly learning optical flow from images but their performances are below the state of the art therein. Ilg et al. [14] stack several FlowNetC and FlowNetS networks into a large model, FlowNet2, which achieves state-of-the-art results and runs faster than traditional methods. However, because of the large size of its model, it always takes dozens of days to train the whole model and it's not feasible to be implemented on mobile and embedded devices. Sun et al. [28] design a compact but effective CNN model, PWC-Net, according to well-established principles: pyramidal processing, warping, and the use of a cost volume, which reduces the size of CNN model and achieves better results on existing benchmarks. There are also works focusing on the robustness of optical flow, such as Robust Flow [16] and RainFlow [17]. They try to keep optical flow accuracy in rainy scenes by introducing residue channels and veiling-invariant features. However, none of these methods pays attention to the performance of optical flow in the dark and current optical flow datasets lack real low-light data.

Synthetic Low-light Dataset. FlyingChairs [10] is a synthetic optical flow dataset containing about 22k image pairs of chairs superimposed on random background images from Flickr. [22] proposes the FlyingThings3D dataset which can be seen as a three-dimensional version of the FlyingChairs. We decide to adopt the FlyingChairs dataset and synthesize low-light effect on it. Many approaches have been proposed to synthesize low-light and noisy image datasets. [19, 18] uses additive Gaussian noise and Gamma correction to simulate low-light effects. [15] simulates the degradation and noise transformation performed by camera pipelines. Very related to our work is [2, 23] in which they add read and shot noise on the raw sensor data, because 8-bit jpeg images are easy to lose information when turned dark, and we aim to derive optical flow directly from dark raw images. In addition, we conduct our noise analysis on real raw images of various brightness, while existing noise analysis is mostly on bright noisy images caused by various ISO.

Low-light Imaging. A variety of techniques have been developed for low-light imaging. Infrared information is commonly used in low-light detection [21, 20], but infrared sensors are not widely equipped, so our experiments focus on RGB image enhancing. Non-local means (NLM) [3] and BM3D [9] are denoising methods based on finding similar patches around every pixel, then average all the patches and replace the pixel with the result. Recently, Chen et al. [6, 5] propose a learning-based method to enhance dark images. They take raw data as input to a modified U-Net [25]. However, even if some of the methods could achieve good perception, it's unavoidable to lose information while post-processing and information like brightness consistency is critical to optical flow estimation.

3. Various Brightness Optical Flow Dataset

There are many benchmarking datasets about low-light and optical flow, such as SID [6] and SIDD [1] that provide low-light or noisy images with their corresponding bright clean images, and KITTI [12] and Sintel [4] that provide image pairs with their corresponding optical flow. However, to our knowledge, no existing dataset is able to benchmark the low-light performance of an optical flow model, so we combine the two concepts above and collect a Various Brightness Optical Flow (VBOF) dataset, which contains 598 raw images of various brightness with 100 corresponding optical flow reference.

The images are taken by 3 cameras: Sony A6000 (102 images), Canon EOS M6 (297 images), Fujifilm XT2 (199 images), in which Canon and Sony have Bayer-pattern sensors while Fujifilm has X-Trans sensor. We mount the camera on a tripod, control it remotely to avoid vibration and set various exposure time to get various exposure photos of the same movement.

Various exposure image pairs share the same optical flow. We took the raw images in about 10 different exposures before and after we move the objects, then we use the bright sharp image pair to calculate the optical flow with [14] to serve as a reference for the dark noisy pairs. Some samples are shown in Figure 2.

Besides the capability to evaluate the brightness robustness of an optical flow model, VBOF also has other advantages. First, we choose to present our data in 14-bit raw format so that we can use various methods to demosaic and enhance images with less information loss. Second, since the VBOF dataset is collected using 3 cameras from different manufacturers and raw format is unique to every camera model, our dataset is able to evaluate the generalization ability for different cameras. Third, the VBOF dataset consists of scenes both indoor and outdoor, with various lighting effects, so it comprehensively reflects scenes in real life.

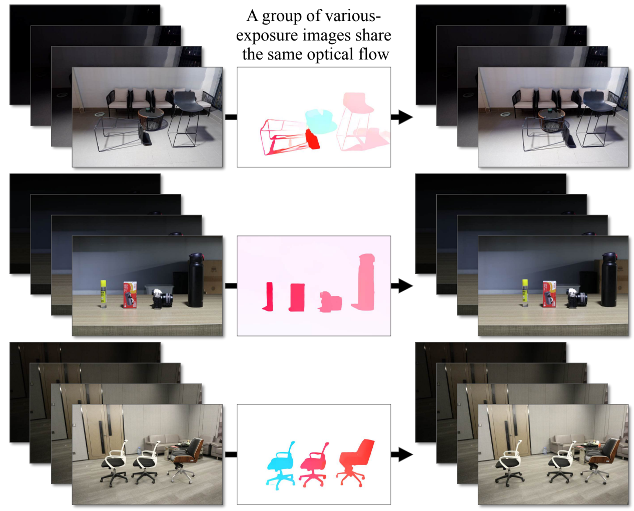


Figure 2. Three samples in our Various Brightness Optical Flow (VBOF) dataset: raw images of moving objects with a large range of brightness with reference optical flow.

4. Method

4.1. End-to-end Optical Flow

We design an end-to-end pipeline to deal with low-light optical flow (showed in Figure 3). The core of our solution is to synthesize low-light raw effects on bright RGB images in the original dataset for training optical flow models and directly use raw as input when testing, instead of enhancing the low-light input before estimation using models trained on the original dataset. We get the inspiration from [27, 26], in which they achieve great image segmentation results on foggy input.

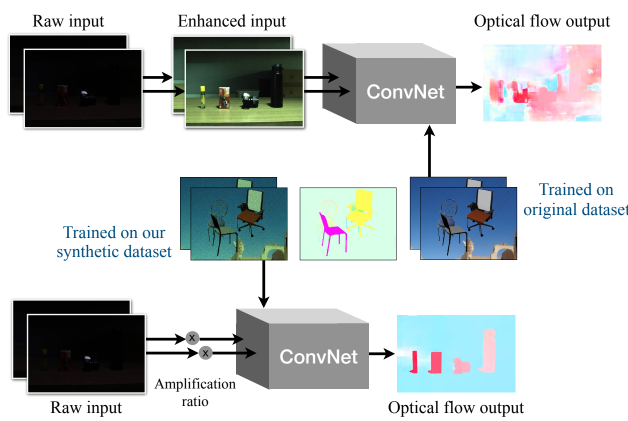
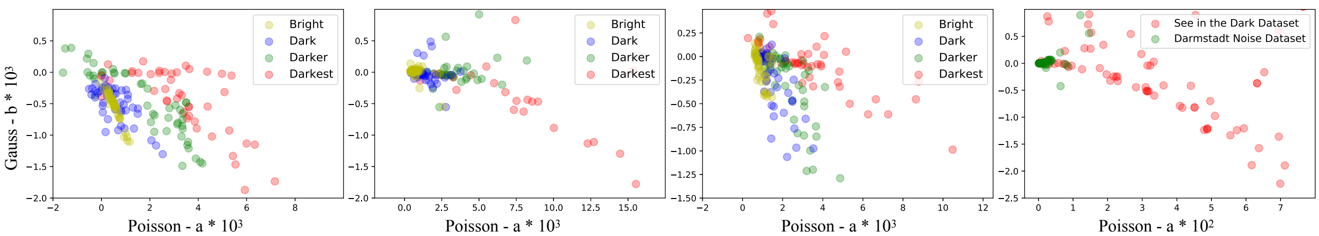


Figure 3. Above: two-step pipeline using existing methods for optical flow in the dark. Below: our pipeline for optical flow in the dark.



(a) Canon (b) Sony (c) Fujifilm (d) SID vs DND

Figure 4. Distribution of Gauss-related parameter a and Poisson-related parameter b analyzed on Canon, Sony and Fujifilm camera of various brightness in (a), (b) and (c) respectively, and on the See in the Dark dataset [6] and the bright Darmstadt Noise Dataset [24] in (d).

4.2. Noise Analysis

Unlike the noise in a processed image which may have complex characteristics because of compression and other nonlinear operations, the noise in raw sensor data is better understood. It generally comes from two sources: photon shot noise and thermal read noise. The number of photons arriving the sensor varies from exposure to exposure and from pixel to pixel, which is governed by the Poisson distribution:

$$P(k_{ph}, \lambda_{br}) = \frac{\lambda_{br}^{k_{ph}} * e^{-\lambda_{br}}}{k_{ph}!} \tag{1}$$

where λ_{br} is the expected number of occurrences which rises as the image gets brighter, and k_{ph} is the number of photons arriving. Various k_{ph} of different pixels on the sensor would lead to unavoidable photon shot noise. And the other main noise source is thermal noise in readout circuitry, which doesn't have a fixed pattern and the only solution to this is cooling, so we simulate the thermal read noise using Gaussian distribution.

Considering the two types of noise and that we want to simulate the image noise of low-light data, we carefully collect a Various Brightness Raw (VBR) dataset, setting the brightness as the only variable. In order to reveal the changing noise pattern as the brightness descends, we fix the exposure time and ISO to get invariant thermal read noise and only change the size of the aperture to get images of different brightness. We also repeat the process on different cameras to get general results. Finally, we get 1200 raw images of a large range of brightness both indoor and outdoor, in the day and the night environments.

To analyze the noise model, we employ the noise estimation method in [11] to estimate the Poisson-related parameter a and the Gaussian-related parameter b. Specifically, we assume that the noisy image z can be decomposed by the following formula:

$$z(x) = y(x) + \eta_p(y(x)) + \eta_g(x) \tag{2}$$

where $x \in X \subset N^2$ is the pixel position, y is the original noise-free image, and

$$(y(x) + \eta_p(y(x)))/a \sim \mathcal{P}(y(x)/a),\ \eta_g(x) \sim \mathcal{N}(0, b) \tag{3}$$

in which $\mathcal{P}$ and $\mathcal{N}$ denote Poisson and Gaussian distribution respectively. We approximate these together as a single heteroscedastic Gaussian:

$$z(x) \sim \mathcal{N}(\mu = x, \sigma^2 = ax + b) \tag{4}$$

The analysis result is showed in Figure 4, where we can see that generally the noise level rises as the brightness descends and that in low-intensity dark images, a and b are more tightly coupled and Poisson-related parameter becomes relatively larger than Gaussian-related parameter, which means the noise is more signal-dependent. It's also obvious that the noise parameters analyzed on the bright Darmstadt dataset [24] which is used in [2] are quite different from the ones we analyze on our low-light dataset and the See in the Dark dataset [6].

4.3. Synthesize New Optical Flow Dataset

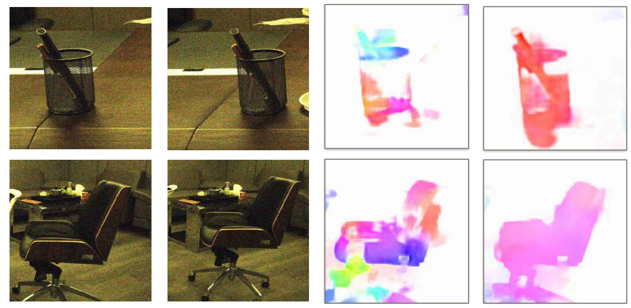
In order to generate synthetic noisy raw images from clean images, we first randomly invert the gamma correction of RGB channels respectively to simulate the uncorrected light effects and white balance on raw images, then we randomly sample shot and read parameter a and b from ranges that match what we observe in real data and add noise by sampling from the distribution of Eq. 4. We choose to perform our operations based on the FlyingChairs [10] dataset after normalizing the 8-bit images in it to [0-1]. Note that we don't reduce the brightness in the 8-bit training data to preserve information. To match the training input, when testing we scaled the low-light 14-bit raw input to a normal brightness since raw format can keep the weak signals.

Specifically, we sample the noise parameters from our VBR dataset, VBOF dataset, and SID [6] dataset, so that the parameters we sample from would cover various brightness and ISO. Also, we deliberately arrange our synthetic dataset to have an average distribution of noise levels. The final training set is called FlyingChairs-DarkNoise (FCDN).

Finally, we analyze the general variance of the real raw data and our synthetic data using the method in [7] showed in the table 1, in which the noise level is defined by $variance * 255$ of Normal distribution.

Datasets	Noise Level ($variance * 255$)			
	$0-10$	$10-20$	$20-30$	>30
VBR-bright	266	34	0	0
VBR-dark	42	248	10	0
VBR-darker	0	46	242	12
VBR-darkest	0	32	202	66
FlyingChairs	45744	0	0	0
FCDN	17021	14214	14166	343

Table 1. Statistics of images numbers of various levels of noise on different datasets analyzed by the method in [7]. VBR is the dataset we collect of Various Brightness Raw images. FlyingChairs is from [10] and FlyingChairs-DarkNoise (FCDN) is the one we synthesize.



(a) Frame 1 (b) Frame 2 (c) Unprocess (d) Ours

Figure 5. A comparison of final results using different training set synthesizing method: "unprocess" [2] and ours. (a,b) are inputs scaled from dark raw images. In our result, the object is painted red and purple meaning moving to the right and upper right, which is the correct output.

5. Experiments

5.1. Setup

To evaluate our method we use our Various Brightness Optical Flow dataset since it's the only dataset that is able to benchmark brightness robustness for optical flow models so far. The performance is evaluated by End-Point Error (EPE), which is the Euclidean norm of the difference between the estimated optical flow vector (V_{est}) and a reference optical flow vector (V_{rf}). V_{est} and V_{rf} have the shape of $(H, W, 2)$ where H and W represent the size of the input image, and 2 represents a 2-dimension vector for each pixel, with its direction and length indicating the direction and speed of the movement in that pixel.

Network Choice. We choose to use FlowNetC, FlowNetS [10] and PWC-Net [28] to evaluate our method. The reason is that they can represent the mainstream two kinds of architectures of optical flow networks (U-Net and spatial pyramid) and many other networks are based on them, such as FlowNet2 [14] which stacks several FlowNetC and FlowNetS networks together. And PWC-Net [28] is one of the state-of-the-art optical flow networks. The three networks are all compact but effective, and they can all be trained in a reasonable amount of time.

Training Set Synthesis. We want to use raw images as input to avoid information loss when testing, so we need to generate raw features on RGB images in the original training dataset, which is the core of our solution. Brooks et al. [2] propose a method to "unprocess" RGB images to RAW for denoising purpose, which includes adding noise, inverting tone mapping, gamma decompression, inverting white balance and digital gain, etc. We try to synthesize a "raw" FlyingChairs dataset using their method but after training and evaluating, we find the final optical flow result turns out to be inaccurate compared to our solution (See Figure 5). We think the main reason is that we analyze the noise and synthesize it based on multiple datasets covering a large range of exposure and ISO, while their analysis is on the DND dataset [24] which is relatively limited (See (d) in Figure 4) and other operations like compression tend to destroy signal in 8-bit images.

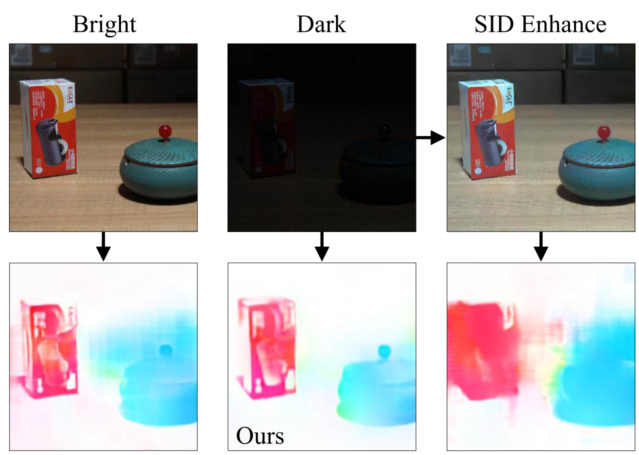


Figure 6. Different inputs of the same movement and optical flow. Our result produced directly from dark input is better than the one produced from the SID enhanced image [6], and even better than the one produced from the original bright image using PWC-Net author-trained model [28].

Evaluating Dataset	Model	Training Set	End-Point Error of Various Exposure Input (0: brightest, 6: darkest)						
			Exp-0	Exp-1	Exp-2	Exp-3	Exp-4	Exp-5	Exp-6
VBOF	FlowNetC	FlyingChairs	5.78	6.21	6.98	7.98	9.48	12.21	14.73
		FCDN	5.76	5.78	5.78	5.86	5.98	6.11	6.55
	FlowNetS	FlyingChairs	11.15	11.34	11.60	11.57	10.77	10.09	9.76
		FCDN	5.46	5.53	5.51	5.53	5.61	5.67	6.26
	PWC-Net	FlyingChairs	7.13	7.04	7.08	7.52	7.58	9.21	10.88
		FCDN	**5.31**	**5.32**	**5.29**	**5.34**	**5.37**	5.63	6.09
VBOF + NLM	FlowNetC	FlyingChairs	5.34	5.39	5.78	7.02	8.96	11.36	13.69
		FCDN	5.89	5.95	5.92	5.92	6.04	6.12	6.55
	FlowNetS	FlyingChairs	11.38	11.81	12.69	11.64	10.44	9.73	9.55
		FCDN	5.80	5.93	5.98	5.83	5.72	5.64	6.25
	PWC-Net	FlyingChairs	6.14	6.15	7.12	7.12	7.31	8.59	10.19
		FCDN	5.50	5.57	5.51	5.50	5.48	**5.57**	**6.04**
VBOF (only Sony part) + SID (Sony model)	FlowNetC	FlyingChairs	5.61	5.75	5.75	5.97	6.40	6.40	6.81
		FCDN	4.74	4.77	4.84	4.87	4.93	5.28	5.13
	FlowNetS	FlyingChairs	9.28	10.05	10.31	10.65	11.27	12.43	11.83
		FCDN	4.47	4.59	4.76	4.78	5.24	6.13	6.22
	PWC-Net	FlyingChairs	**4.27**	4.41	4.58	4.99	5.35	5.58	5.48
		FCDN	4.31	**4.27**	**4.25**	**4.37**	**4.30**	**4.41**	**4.60**

Table 2. Performance of our solution compared to existing solutions. The evaluation is done on our Various Brightness Optical Flow (VBOF) dataset and the processed VBOF dataset using Non-local Means [3] and Learning to See in the Dark (SID) [6]. Because the SID method is camera-sensitive, relevant experiments are only done on the Sony part of our VBOF dataset. We choose FlowNetC and FlowNetS [10] and PWC-Net [28] as our optical flow model and they are trained on the FlyingChairs dataset [10] and our FlyingChairs-DarkNoise (FCDN) dataset. The VBOF dataset is separated by multiple brightness levels, we select 7 of them listed as 'Exp-0' to 'Exp-6'. From the table we can see that models trained on our dataset always get a performance improvement almost on every brightness level and the accuracy is much more stable while the brightness is changing.

5.2. "Bad" Enhancement

The direct solution to get optical flow from dark images is to apply an image-enhancing method before optical flow estimation. Scaling is the simplest way to enhance a dark image and it is a reversible process with little information loss. However, because the scaled raw images suffer from serious noise and color aberration, it unsurprisingly leads to poor optical flow results using existing methods. We then try advanced methods — a traditional enhancing method and a learning-based enhancing method.

We present the result of Non-local Means (NLM) [3] as a representative of traditional enhancing methods. It searches similar patches and averages them together to remove the noise. From Figure 7 input block, we can see NLM doesn't correct the color aberration but it produces a relatively stable denoising result. From table 2, we can see that followed by an optical flow model trained on existing FlyingChairs dataset [10], NLM doesn't solve the problem that optical flow models' performance gets worse on low-light images, which is much worse than our result produced from VBOF dataset without NLM processing.

We present the result of Learning to See in the Dark (SID) [6] as a representative of learning-based enhancing methods. Chen et al. [6] propose to use U-Net [25] to enhance dark raw images, and they provide their model trained on Sony raw images, so we also use the Sony part of our VBOF dataset to test the method, and we get visually great images using the provided model. (See Figure 6,7) SID does a good job on white balance and the result images are noiseless, but if zooming in to take a closer look, one will find the signal in the image is unstable especially for the extremely dark images. Followed by an optical flow model trained on the existing FlyingChairs dataset [10], SID also doesn't solve the problem of optical flow in the dark. Sometimes SID may lead to worse optical flow results showed in Figure 7.

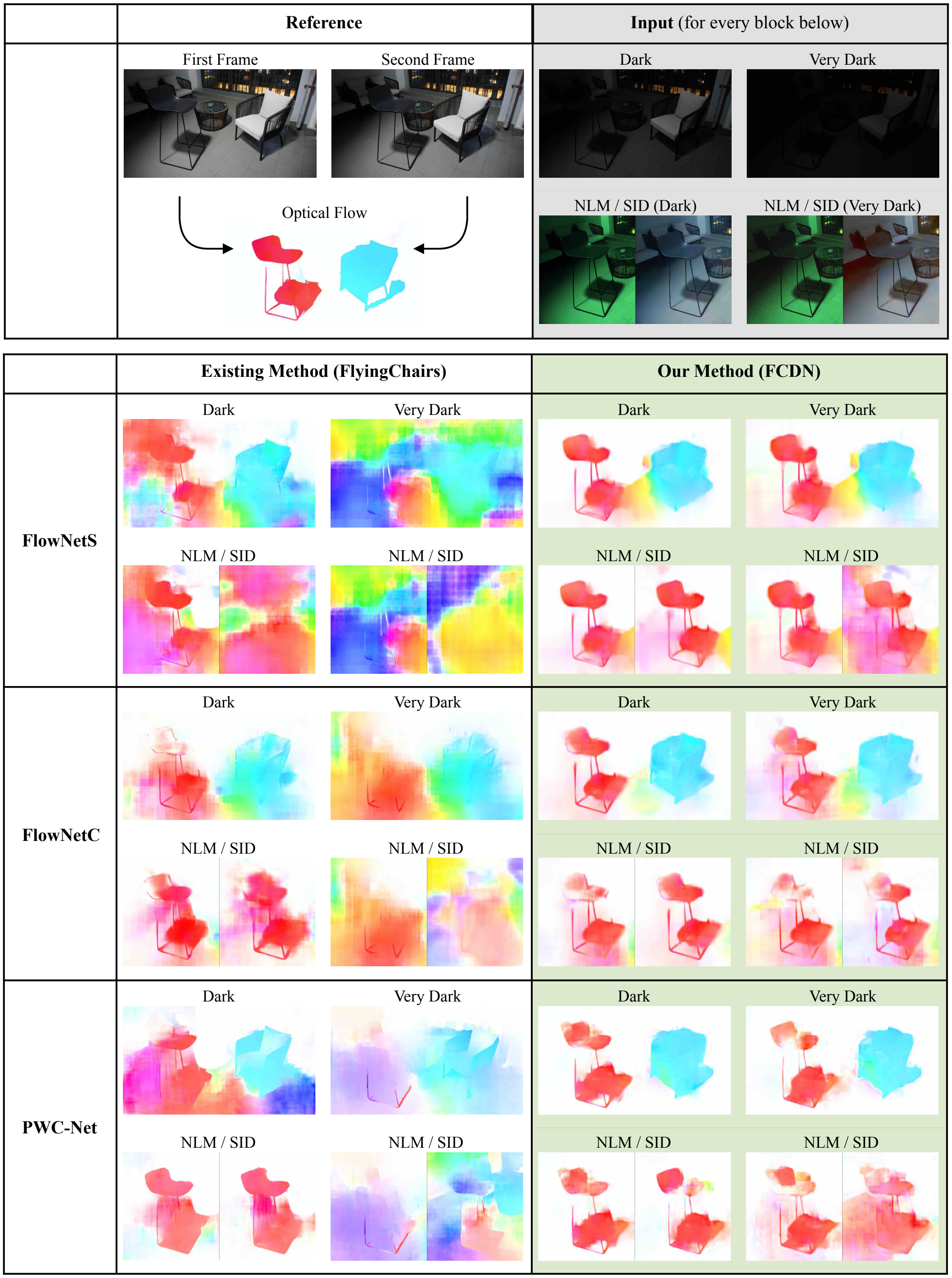


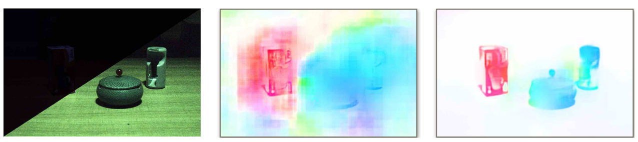
Figure 7. A detailed evaluation of existing methods and our method. The upper right block serves as the input for all the blocks below, where we evaluate the performance of deriving optical flow from dark images and the enhanced images [3, 6] with various optical flow models [10, 28]. The upper left block serves as a good reference with bright input images and optical flow.

5.3. Performance Details

From Table 2 and Figure 7, we can see that for every optical flow model, the ones trained on our FCDN dataset always show better stability when the brightness descends and they have better accuracy dealing with extreme low-light image inputs.

(a) Sony outdoor ISO 1000

(b) Canon indoor ISO 3200



(c) Fujifilm outdoor ISO 2500

Figure 8. Generalization performance of our solution (right) compared to existing methods (middle) evaluated with PWC-Net[28] on images in our VBOF dataset (left) of various cameras, ISO, indoor and outdoor.

There are some other advantages of our solution that are worth pointing out. First, given the same optical flow network and the same light condition, our method's performance on the underexposed input is even better than the existing method's performance on well-exposed input (See Figure 6, Table 2). The reason is that a well-exposed image captured with long exposure or high ISO has different noise distribution from that of a bright one captured in a well-illuminated condition, and the network trained on the latter might not generalize well to the former bright image. Through realistic noise simulation, the network trained on our synthesized data is able to deal with various noise at different brightness levels. Second, from Figure 7,8, we can see that our method can generalize on different camera models in various environments and it's effective for multiple optical flow models.

6. Discussion

Although PWC-Net [28] achieves a higher accuracy on many of optical flow benchmark datasets using a 17 times smaller model than the FlowNet2 model [14], large models like FlowNet2 still performs better on complex real data like the well-exposed images in our VBOF dataset. From this point of view and our statistics, besides achieving a higher performance on low-light images, training a compact model like PWC-Net on our FCDN dataset can also help such small optical flow models learn to deal with complex situations in real images of any exposure more accurately.

There are also works published achieving relevant goals in the semantic image segmentation area by adapting models from daytime to nighttime [8] and from clear weather conditions to foggy conditions [26], which bears resemblance to works from the broad field of transfer learning. They mix synthetic data and real data, then gradually adapt a model from clean easy training data to corrupted hard training data, such as from clear weather to light synthetic fog, and finally to dense real fog in multiple steps. We have already proved that directly training optical flow models on our synthetic low-light data is effective, and we believe that training models with our data in a manner of gradually model adaption may also lead to promising results, which can be tried in future work.

7. Conclusion

In this paper, we have presented a data-driven solution to improve optical flow accuracy especially in low-light environments. By synthesizing training data based on the noise model we analyzed on raw images collected in various brightness conditions, we succeed in training optical flow models that outperform the state-of-the-art on the real low-light optical flow dataset — Various Brightness Optical Flow (VBOF) dataset that we collect. VBOF consists of 598 raw images of various brightness with corresponding reference optical flow, aiming to benchmark the brightness robustness of optical flow models. We believe the proposed method, the noise analysis, and the VBOF dataset will be very useful for optical flow tasks in real scenes.

References

[1] Abdelrahman Abdelhamed, Stephen Lin, and Michael S. Brown. A high-quality denoising dataset for smartphone cameras. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018. 3

[2] Tim Brooks, Ben Mildenhall, Tianfan Xue, Jiawen Chen, Dillon Sharlet, and Jonathan T Barron. Unprocessing images for learned raw denoising. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 11036–11045, 2019. 2, 4, 5

[3] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. Non-local means denoising. *Image Processing On Line*, 1:208–212, 2011. 2, 3, 6, 7

[4] Daniel J Butler, Jonas Wulff, Garrett B Stanley, and Michael J Black. A naturalistic open source movie for optical flow evaluation. In *European conference on computer vision*, pages 611–625. Springer, 2012. 2, 3

[5] Chen Chen, Qifeng Chen, Minh Do, and Vladlen Koltun. Seeing motion in the dark. 2019. 3

[6] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun.

Learning to see in the dark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3291–3300, 2018. 1, 2, 3, 4, 5, 6, 7

[7] Guangyong Chen, Fengyuan Zhu, and Pheng Ann Heng. An efficient statistical method for image noise level estimation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 477–485, 2015. 5

[8] Dengxin Dai and Luc Van Gool. Dark model adaptation: Semantic image segmentation from daytime to nighttime. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 3819–3824. IEEE, 2018. 8

[9] Aram Danielyan, Vladimir Katkovnik, and Karen Egiazarian. Bm3d frames and variational image deblurring. *IEEE Transactions on Image Processing*, 21(4):1715–1728, 2011. 2, 3

[10] Alexey Dosovitskiy, Philipp Fischer, Eddy Ilg, Philip Hausser, Caner Hazirbas, Vladimir Golkov, Patrick Van Der Smagt, Daniel Cremers, and Thomas Brox. Flownet: Learning optical flow with convolutional networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2758–2766, 2015. 2, 4, 5, 6, 7

[11] Alessandro Foi, Mejdi Trimeche, Vladimir Katkovnik, and Karen Egiazarian. Practical poissonian-gaussian noise modeling and fitting for single-image raw-data. *IEEE Transactions on Image Processing*, 17(10):1737–1754, 2008. 4

[12] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. Vision meets robotics: The kitti dataset. *The International Journal of Robotics Research*, 32(11):1231–1237, 2013. 2, 3

[13] Berthold KP Horn and Brian G Schunck. Determining optical flow. *Artificial intelligence*, 17(1-3):185–203, 1981. 1

[14] Eddy Ilg, Nikolaus Mayer, Tonmoy Saikia, Margret Keuper, Alexey Dosovitskiy, and Thomas Brox. Flownet 2.0: Evolution of optical flow estimation with deep networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2462–2470, 2017. 1, 2, 3, 5, 8

[15] Ronnachai Jaroensri, Camille Biscarrat, Miika Aittala, and Frédo Durand. Generating training data for denoising real rgb images via camera pipeline simulation. *arXiv preprint arXiv:1904.08825*, 2019. 2

[16] Ruoteng Li, Robby T Tan, and Loong-Fah Cheong. Robust optical flow in rainy scenes. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 288–304, 2018. 2

[17] Ruoteng Li, Robby T Tan, Loong-Fah Cheong, Angelica I Aviles-Rivero, Qingnan Fan, and Carola-Bibiane Schonlieb. Rainflow: Optical flow under rain streaks and rain veiling effect. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 7304–7313, 2019. 2

[18] Kin Gwn Lore, Adedotun Akintayo, and Soumik Sarkar. Llnet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition*, 61:650–662, 2017. 2

[19] Feifan Lv, Feng Lu, Jianhua Wu, and Chongsoon Lim. Mbllen: Low-light image/video enhancement using cnns. In *BMVC*, page 220, 2018. 2

[20] Feifan Lv, Yinqiang Zheng, Yicheng Li, and Feng Lu. An integrated enhancement solution for 24-hour colorful imaging. In *Proceedings of the AAAI conference on artificial intelligence*, 2020. 2, 3

[21] Feifan Lv, Yinqiang Zheng, Bohan Zhang, and Feng Lu. Turn a silicon camera into an ingaas camera. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5987–5995, 2019. 3

[22] Nikolaus Mayer, Eddy Ilg, Philipp Hausser, Philipp Fischer, Daniel Cremers, Alexey Dosovitskiy, and Thomas Brox. A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4040–4048, 2016. 2

[23] Ben Mildenhall, Jonathan T Barron, Jiawen Chen, Dillon Sharlet, Ren Ng, and Robert Carroll. Burst denoising with kernel prediction networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2502–2510, 2018. 2

[24] Tobias Plotz and Stefan Roth. Benchmarking denoising algorithms with real photographs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1586–1595, 2017. 2, 4, 5

[25] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015. 3, 6

[26] Christos Sakaridis, Dengxin Dai, Simon Hecker, and Luc Van Gool. Model adaptation with synthetic and real data for semantic dense foggy scene understanding. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 687–704, 2018. 3, 8

[27] Christos Sakaridis, Dengxin Dai, and Luc Van Gool. Semantic foggy scene understanding with synthetic data. *International Journal of Computer Vision*, 126(9):973–992, 2018. 3

[28] Deqing Sun, Xiaodong Yang, Ming-Yu Liu, and Jan Kautz. Pwc-net: Cnns for optical flow using pyramid, warping, and cost volume. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8934–8943, 2018. 1, 2, 5, 6, 7, 8