International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)

Development of Naïve technique for Removing Fences, Reflections and Raindrops from image Pattern Recognition Techniques

Arvind Singh[1], Shailesh Kumar Singh[2]

Research Scholar[1], AP[2], SHEAT COLLEGE OF ENGINEERING, VARANASI 384

Article Info

Volume 9, Issue 6

Page Number : 48-55

Publication Issue

November-December-2022

Article History

Accepted : 01 Nov 2022

Published : 04 Nov 2022

Abstract – Mobile, Smartphone, Tablets, and other devices with built-in cameras are now widely available at affordable prices, allowing regular people to capture and share life's most important moments with the world. However, a novice photographer may not be pleased with the results after posting them online. It's possible that the photographer's subject of interest has been blocked off or walled off in some way. The authors write, "We provide a learning-based technique for removing unwanted barriers from a rapid succession of photographs obtained by a moving camera, such as window reflections, fence occlusions, or rainfall. It takes into account the relative velocity of the foreground and background. Inaccuracies in the flow estimate and brittle assumptions like brightness consistency may be accounted for thanks to the learning-based layer reconstruction. We show how well training on synthetic data generalises to real-world photographs. Our research into many challenging reflection and fence removal scenarios has shown encouraging results, demonstrating the efficacy of the proposed approach.

Keywords – Deep Convolutional Neural Network, Fence Removal, Flow Estimation, Dense Optical Flow

I. INTRODUCTION

We often have to snap pictures in less-than-ideal lighting, with obstacles like windows or other objects in the way. Reflections of inside items, for instance, may ruin a photograph taken via a window of the outside. Shooting through a fence or enclosure may be necessary if we want to photograph zoo animals (Fig. 1.1). Changing the camera's location or plane of focus isn't always adequate to eliminate such visual obstacles, and current computational methods still aren't compelling when removing them from photographs.

However, more advanced alternatives, like polarised lenses (to eliminate reflections), are out of reach for the average consumer. In this research, we provide a robust algorithm that enables a user to capture pictures through barriers like windows and fences while still capturing the area of interest as if the obstacles weren't there. When using our technique, users need only create slight camera motion throughout the imaging process; all other processing is handled mechanically.

Therefore, we tell the photographer to capture a series of images while gently moving the camera, much as they would do when capturing a panorama, rather than just one single shot. Our system then combines the space-time information to generate two images: one of the backdrop and one of the reflected or occluded content, based on changes in the movements of the layers caused by visual parallax. Although there are certain limitations to the scene that our arrangement imposes (for example, maintaining a relatively constant position while taking pictures), we found that frequent shooting settings are amenable to it.

Motion parallax has been used before in layer decomposition. Instead, the fundamental contribution of our study is a more stable and accurate technique for motion estimation when obstacles are present. In our method, the first step is calculating sparse motion fields on the edges of pictures after receiving an input picture sequence. This is followed by a coarse-to-fine estimation of backdrop and obstacle layers before we apply interpolation to turn the sparse edge flows into more dense motion fields.

We also demonstrate that a single framework can deal with reflections and physical occlusions. At first appearance, reflections and occlusions may seem quite different, and various approaches have been used to deal with each. While previous works have taken separate approaches to these issues, the method presented here takes a holistic view of the situation. Our system comprises mostly shared modules for these two issues, and with just minor adjustments, it outperforms prior methods that specifically addressed either subproblem. In this study, we consider both reflecting and occluding obstructions and explicitly select which kind is present in the picture to fine-tune the algorithm.

We tested our technology in various realistic and natural settings. Instead of mimicking obstacles artificially via the blending or composting of pictures (as was widely done in prior work), we conduct controlled trials to capture real situations with ground truth decompositions for quantitative assessment. Our technique is automated, compatible with most smartphone cameras, and needs just free-form camera movement from the user to scan the area. As a result of our research, we have determined that five photographs were taken along a narrow..

When photographing via reflecting surfaces or obstructing objects, it might be challenging to get clear shots since the resulting photos will include both the subjects of interest and the barriers. To enhance the quality of images acquired in such situations or to enable computers to create a proper physical interpretation of the environment.

Automatically deleting undesired reflections or occlusions from a single picture has been the subject of recent research [2, 8, 16, 17, 27, 38, 43, 45]. These techniques use ghosting signals [30] or learning-based approaches [8, 16, 38, 43, 45] to recapture the original natural visuals. While excellent results have been shown, the task of differentiating the clean backdrop from occlusions is inherently ill-posed and sometimes needs a high-level semantic comprehension of the scene to do correctly. For photos that fall beyond the norm, the effectiveness of learning-based approaches drops dramatically.

The central concept is to use the depth difference between the camera and the foreground, backdrop, and occluding items. Accordingly, the motion differences between the two layers may be seen in photos taken with a slightly moving camera [3, 9, 12, 21, 24, 34]. Several methods [1, 3, 6, 9, 12, 21, 24, 26, 31, 34] use these signals to remove reflections or fences from videos. Using a unified computational framework, Xue et al. [42] demonstrate outstanding results on several natural sequences and offer a method for removing obstructions. However, the formulation needs a computationally costly optimization procedure and is highly dependent on either the assumption of constant brightness or the accuracy of motion estimates. Current research [1] investigates model-free approaches by using a general-purpose 3D-CNN to address these concerns. Unfortunately, when applied to real-world input sequences, CNN-based approaches cannot compete with the quality of the results obtained by optimization-based algorithms.

A multi-frame obstacle removal technique that combines the best characteristics of the previous optimization- and learning-based approaches is presented in this thesis. Using an optimization-based strategy, the suggested method uses a coarse-to-fine system, alternating between dense motion estimates and background/obstruction layer reconstruction phases. The mindless motion is precisely described to bring back layer-specific information progressively. Our fusion network is readily transferrable to real-world sequences despite being trained on a lab dataset. Our results demonstrate that the proposed method beats current available algorithms on various demanding sequences and applications."

In contrast to the optimization-based formulation of [26, 42], our model is entirely data-driven and does not assume any classical properties of the scene, such as the scene's brightness being constant, the flow field's accuracy, or the surface's being flat. When these conditions are not met, it might be difficult for traditional methods to rebuild distinct layers of background and foreground. The data-driven strategies can learn from a wide variety of training data and accept erroneous results when these assumptions are not met. In this work, we provide a learning-based approach for recovering clean pictures from a given brief series of photos captured by a moving camera via obstructive features such as (1) windows, (2) fences, or (3) rainfall.

II. RATIONALE FOR THE STUDY

A small number of studies address the issue of improving vision in adverse weather conditions; this thesis generally deals with haze or fog and rain streaks. Several strategies have been developed for raindrop detection. Using principal component analysis, Kurihara et al. [12] study the form of raindrops and then try to match an area in the test picture to the raindrops they've learned. A strategy proposed by Roser and Geiger [17] is to compare a synthetically manufactured raindrop to a spot that could have a raindrop. Initially, it is assumed that the synthetic raindrops are spherical in cross-section, and subsequently, it is supposed that they are oblate in cross-section [18]. Since raindrops may come in various forms and sizes, these assumptions could work in certain situations, but they can't be applied universally. Reflection and occlusion removal are two image processing challenges that have been studied in the past using a variety of approaches. Here, we conduct a literature review of the relevant research in these two fields. Reconstructing the backdrop and obstacle layers allows us to fine-tune optical flows to the final layer. The whole flow of our process is shown in Figure 2. Several layer decomposition issues, including reflection/obstruction/fence/rain removal, are amenable to our unified architecture. Without sacrificing generality, we present our approach through the reflection removal job. In the following sections, we discuss the three modules' specifics.

III. PROBLEM STATEMENT

Is there a technique of image processing which can remove undesirable reflections, Rain drop, fancy on a sean from a single digital image.

IV. RESEARCH OBJECTIVES

To answer our issue statement, this research has the following objectives: To develop an algorithm to remove undesirable reflections, Rain drop, fancy on a sean from a single digital image.

V. PROPOSED METHODOLOGY

A. Overview

It's not easy to separate the foreground from the backdrop in a video. Reconstructing the backdrop and obstacle layers allows us to fine-tune optical flows to the final layer. The whole flow of our process is shown in Figure 2. Several layer decomposition issues, including reflection/obstruction/fence/rain removal, are amenable to our unified architecture. Without sacrificing generality, we present our approach through the reflection removal job. In the following sections, we discuss the three modules' specifics.

B. Initial Flow Decomposition

As a foundational step in our approach, we begin by making predictions about the flow at the scene's coarsest level (l = 0), including the backdrop and reflection layers. Instead of guessing at complex flow fields, we suggest learning a single uniform motion vector for each layer. Two modules make up our first flow decomposition network: A layer flow estimator, and A feature extractor.

C. *Background Reconstruction*

Background and reflection reconstruction have the same purposes, but the two layers' properties are quite different. In many shots, for instance, the backdrop layers will seem more prominent than they are. In contrast, reflection layers are often hazy and subdued. Therefore, we separately train a network to rebuild the backdrop and a separate network to reconstruct the reflection layer. Though they have a similar design, the parameters of these two networks are different. The network for reconstructing the reflection layer is identical to the one described below; we only explain the former here.

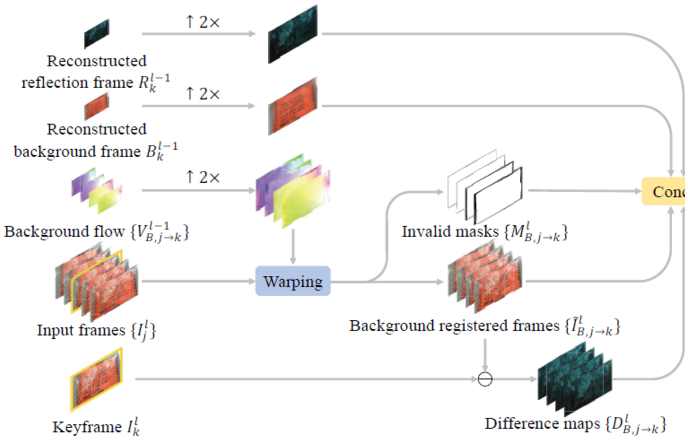

Fig 1: Overview of layer reconstruction module.

While backdrop and reflection reconstruction serve similar purposes, their properties couldn't be more different. Consequently, we replicate the reflection layer by separately training a network and then utilizing the latter to reconstruct the backdrop. These two networks have a similar topology, but their parameters are unique. We only cover the former as the network for reconstructing the reflection layer is the same as the one described below.

D. Network Training

We use a two-stage training process to increase consistency throughout training. The "initial flow decomposition network" is trained in the first step with the following loss:

$$\mathcal{L}_{\text{dec}} = \sum_{k=1}^{T} \sum_{j=1, j\neq k}^{T} \|V_{B,j\to k}^{0} - \text{PWC}(\hat{B}_j, \hat{B}_k)\downarrow^{2^L}\|_1 + \|V_{R,j\to k}^{0} - \text{PWC}(\hat{R}_j, \hat{R}_k)\downarrow^{2^L}\|_1 ,$$

$$\mathcal{L}_{\text{img}} = \frac{1}{T\times L}\sum_{t=1}^{T}\sum_{l=0}^{L}(\|\hat{B}_t^l - B_t^l\|_1 + \|\hat{R}_t^l - R_t^l\|_1), \quad (7)$$

and a gradient loss:

$$\mathcal{L}_{\text{grad}} = \frac{1}{T\times L}\sum_{t=1}^{T}\sum_{l=0}^{L}(\|\nabla\hat{B}_t^l - \nabla B_t^l\|_1 + \|\nabla\hat{R}_t^l - \nabla R_t^l\|_1), \quad (8)$$

$$\mathcal{L} = \mathcal{L}_{\text{img}} + \lambda_{grad}\mathcal{L}_{\text{grad}},$$

E. Synthetic Sequence Generation

To train, we synthesize sequences since acquiring simple lines with ground-truth reflection and backdrop layers is challenging. We randomly choose the backdrop and reflection layers from the 91,000 plus sequences in the Vimeo-90k training set. At first, we apply random homography transformations to bend the sequences. The appendices expand on the topic of synthetic data production.

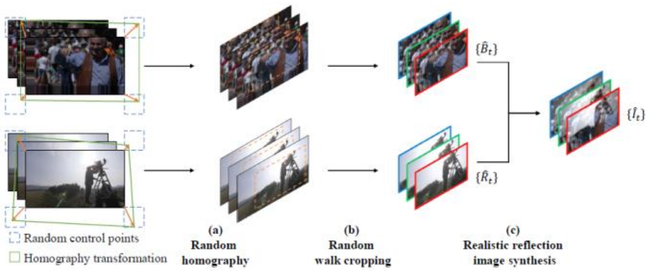

Fig 2: Reflection sequence generation.

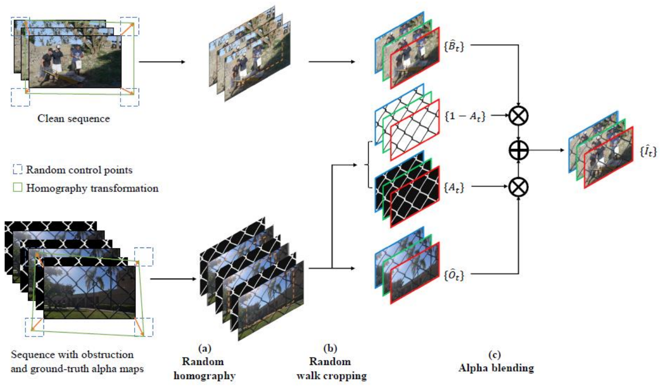

Fig. 3: Obstruction sequence generation.

First, we use randomization to choose a sequence that is unobstructed and one that has a fence or other obstacle. In a manner analogous to the development of the reflection sequence, we apply random homography and random cropping to two series, in addition to the ground-truth alpha maps of the fences or obstacles. After that, we create a new sequence with railings or other barriers using an alpha blending technique.

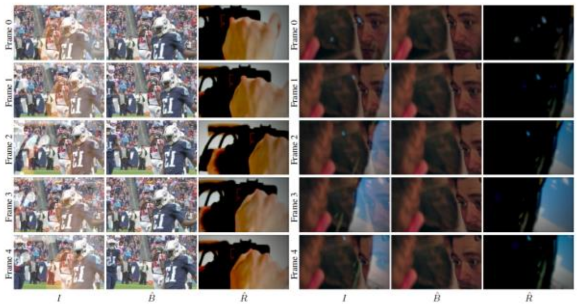

Fig. 4: Training Pairs Generated

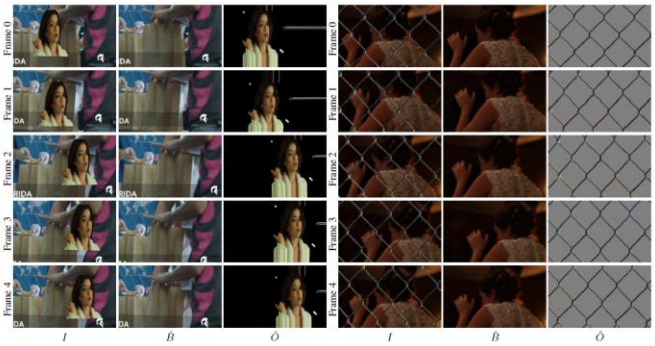

Fig. 5: Training Pairs Generated

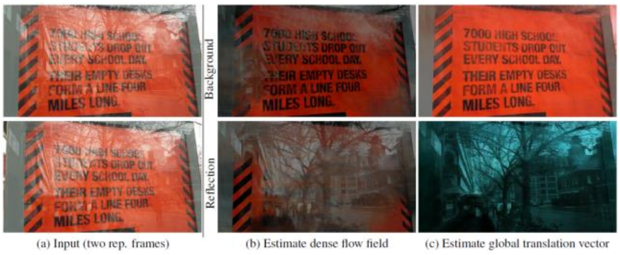

Fig. 6: Initialization with Global Translation Vectors

VI. REFLECTION AND BACKGROUND LAYER RECONSTRUCTION

In Fig. 7, we demonstrate that the model applying a temporal mean or median filter for picture reconstruction does not perform well and frequently creates ghosting artifacts. On the other hand, the suggested image reconstruction network can correct alignment errors induced by the flow estimate in the preceding level and fuses the flow-warped pictures into artifact-free images.

Fig. 7: Image Reconstruction Network

Optimizing for the web rather than losing out on television is shown to be more accurate in noisy predictions in Figure 3.10. Regularizing sparse image gradient priors is one way TV loss aids the network in making smooth predictions.

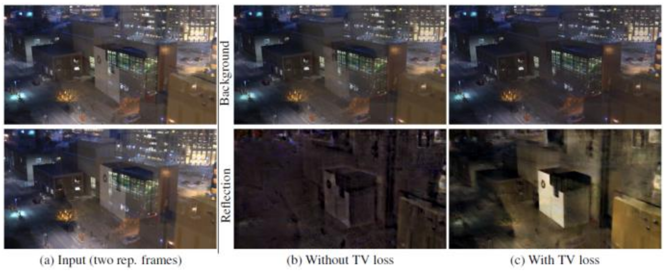

Fig. 8: Online Optimization with TV Loss

VII. EXPERIMENTS & ANALYSIS

We will give the most critical findings in this part, and the supplemental material will contain further discoveries. On a synthetic dataset consisting of one hundred sequences, each of which has five frames in succession, we evaluate the suggested technique compared to other existing methods for removing reflections. We create the findings for the approaches based on a single picture [8, 16, 38, 43, 45] on a frame-by-frame basis. A quantitative comparison of several approaches for removing reflections from synthetic sequences is shown in Table 1.

Table 1: Quantitative Comparison

Method			Background				Reflection			
			PSNR ↑	SSIM ↑	NCC ↑	LMSE ↓	PSNR ↑	SSIM ↑	NCC ↑	LMSE ↓
Single image	CEILNet [8]	CNN-based	20.35	0.7429	0.8547	0.0277	-	-	-	-
	Zhang et al. [45]	CNN-based	19.53	0.7584	0.8526	0.0207	18.69	0.4945	0.6283	0.1108
	BDN [43]	CNN-based	17.08	0.7163	0.7669	0.0288	-	-	-	-
	ERRNet [38]	CNN-based	22.42	0.8192	0.8759	0.0177	-	-	-	-
	Jin et al. [16]	CNN-based	18.65	0.7597	0.7872	0.0218	11.44	0.3607	0.4606	0.1150
Multiple images	Li and Brown [21]	Optimization-based	17.12	0.6367	0.6673	0.0604	7.68	0.2670	0.3490	0.1214
	Guo et al. [12]	Optimization-based	14.58	0.5077	0.5802	0.0694	14.12	0.3150	0.3516	0.1774
	Alayrac et al. [1]	CNN-based	23.62	0.7867	0.9023	0.0200	21.18	0.6320	0.7535	0.1517
	Ours w/o online optim.	CNN-based	**26.57**	**0.8676**	**0.9380**	**0.0125**	**21.42**	**0.6438**	**0.7613**	**0.1008**

We display the keyframe (on the left), the recovered backdrop (in the centre), and the reflection/occluder (on the right) for every sequence (right). As a means of quantitative comparison, we offer the NCC ratings of recovered backgrounds and reflections.

At the beginning of each sequence, we display the keyframe on the left, the recovered backdrop in the centre, and the reflection and occlude states on the right (right). To facilitate quantitative comparisons, we present the NCC scores of the restored backgrounds and reflections.

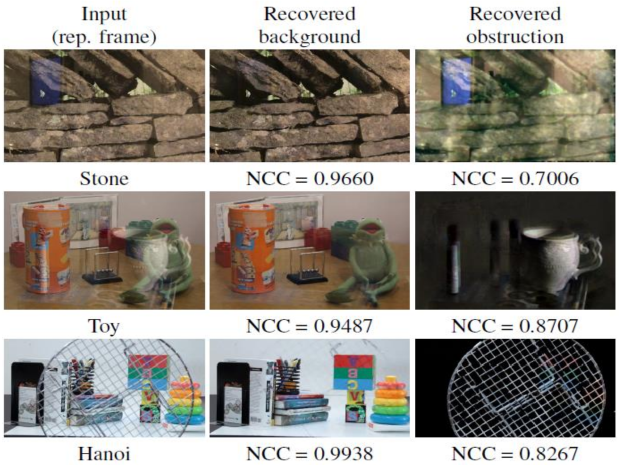

Fig. 9: Quantitative evaluation of controlled sequences

VIII. COMPARISONS AMONG TECHNIQUES

Table 2: Quantitative evaluation of controlled sequences

Method	Stone		Toy		Hanoi	
	B	R	B	R	B	O
Li and Brown [21]	0.9271	0.2423	0.7906	0.6084	-	-
Guo et al. [12]	0.7258	0.1018	0.7701	0.6860	-	-
Xue et al. [42]	**0.9738**	**0.8433**	0.8985	0.7536	0.9921	0.7079
Alayrac et al. [1]	0.9367	0.1633	0.7985	0.5263	-	-
Ours	0.9660	0.7006	**0.9487**	**0.8707**	**0.9938**	**0.8267**

IX. ANALYSIS AND DISCUSSION

We illustrate that uniform flow initialization is important by proving its significance, which demonstrates that it plays a vital role in our strategy. During the training of our model, the following configurations were utilised: deleting the initial flow decomposition network, in which the flows at the coarsest level are set to zero; and forecasting dense flow fields that vary geographically as the initial flows in the model. Both of these configurations were applied in conjunction with one another. In this section, we will investigate a number of significant design choices that were taken for the proposed framework. In addition, we provide an example of a scenario in which our solution is ineffective and propose a time estimate for its implementation.

X. CONCLUSION

Five nearby frames are fed into the technique, and the separation results for the reference frame are generated. Although our technique predicts each reference frame separately, yet provides temporally consistent results for the whole movie. We demonstrate how successfully the suggested technique separates the background and reflection layers while maintaining the input sequences' temporal coherency. In this work, we provide a unique method for removing reflections and obstacles from multiple images. One of our main contributions is using a convolutional neural network to recover flow-warped images' reflection and background layers. Since our method integrates optical flow predictions with coarse-to-fine refining, it can effectively obtain the underlying clean picture from complex real-world sequences. With some tweaks to our design, our technology might be used for a wide variety of purposes, such as removing fences or raindrops. We also demonstrate that the visual quality may be improved by online testing sequences of different settings. Extensive quantitative examination and visual comparisons show that our strategy works well in a variety of settings.

XI. REFERENCES

[1]. Jean-Baptiste Alayrac, Joao Carreira, and Andrew Zisserman. The visual centrifuge: Model-free layered video representations. In CVPR, 2019.

[2]. Nikolaos Arvanitopoulos, Radhakrishna Achanta, and Sabine Susstrunk. Single image reflection suppression. In CVPR, 2017.

[3]. Efrat Be'Ery and Arie Yeredor. Blind separation of superimposed shifted images using parameterized joint diagonalization. TIP, 17(3):340–353, 2008. 1

[4]. Sean Bell, Kavita Bala, and Noah Snavely. Intrinsic images in the wild. ACM TOG, 33(4):159, 2014.

[5]. Yuhua Chen, Cordelia Schmid, and Cristian Sminchisescu. Self-supervised learning with geometric constraints in the monocular video: Connecting flow, depth, and camera. In ICCV, 2019.

[6]. Chen Du, Byeongkeun Kang, Zheng Xu, Ji Dai, and Truong Nguyen. Accurate and efficient video de-fencing using convolutional neural networks and temporal information. In ICME, 2018.

[7]. Elmar Eisemann and Fre´do Durand. Flash photography enhancement via intrinsic relighting. ACM TOG, 23(3):673– 678, 2004.

[8]. Qingnan Fan, Jiaolong Yang, Gang Hua, Baoquan Chen, and David Wipf. A deep generic architecture for single image reflection removal and image smoothing. In ICCV, 2017.
[9]. Kun Gai, Zhenwei Shi, and Changshui Zhang. Blind separation of superimposed images with unknown motions. In CVPR, 2009.
[10]. Kun Gai, Zhenwei Shi, and Changshui Zhang. Blind separation of superimposed moving images using image statistics. TPAMI, 34(1):19–32, 2011.
[11]. Roger Grosse, Micah K Johnson, Edward H Adelson, and William T Freeman. Ground truth dataset and baseline evaluations for intrinsic image algorithms. In ICCV, 2009.
[12]. Xiaojie Guo, Xiaochun Cao, and Yi Ma. Robust separation of reflection from multiple images. In CVPR, 2014.
[13]. Jia-Bin Huang, Sing Bing Kang, Narendra Ahuja, and Johannes Kopf. Temporally coherent completion of dynamic video. ACM TOG, 35(6):196, 2016.
[14]. Shachar Ilan and Ariel Shamir. A survey on data-driven video completion. Computer Graphics Forum, 34(6):60–85, 2015.
[15]. Junho Jeon, Sunghyun Cho, Xin Tong, and Seungyong Lee. Intrinsic image decomposition using structure-texture separation and surface normals. In ECCV, 2014.
[16]. Meiguang Jin, Sabine Su¨sstrunk, and Paolo Favaro. Learning to see through reflections. In ICCP, 2018.
[17]. Sankaraganesh Jonna, Krishna K Nakka, and Rajiv R Sahay. Deep learning-based fence segmentation and removal from an image using a video sequence. In ECCV, 2016.
[18]. Sankaraganesh Jonna, Sukla Satapathy, and Rajiv R Sahay. Stereo image de-fencing using smartphones. In ICASSP, 2017. 2Zdenek Kalal, Krystian Mikolajczyk, and Jiri Matas. Tracking-learning-detection. TPAMI, 34(7):1409–1422, 2011.
[19]. Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In ICLR, 2015.
[20]. Yu Li and Michael S Brown. Exploiting reflection change for automatic reflection removal, in ICCV, 2013.
[21]. Yu Li and Michael S Brown. Single image layer separation using relative smoothness. In CVPR, 2014.
[22]. Zhengqin Li, Mohammad Shafiei, Ravi Ramamoorthi, Kalyan Sunkavalli, and Manmohan Chandraker. Inverse rendering for complex indoor scenes: Shape, spatially-varying lighting, and svbrdf from a single image. In CVPR, 2020.
[23]. Ce Liu, Jenny Yuen, Antonio Torralba, Josef Sivic, and William T Freeman. Sift flow: Dense correspondence across different scenes, in ECCV, 2008.
[24]. Yadong Mu, Wei Liu, and Shuicheng Yan. Video de-fencing. IEEE Transactions on Circuits and Systems for Video Technology, 24(7):1111–1121, 2013.
[25]. Ajay Nandoriya, Mohamed Elgharib, Changil Kim, Mohamed Hefeeda, and Wojciech Matusik. Video reflection removal through Spatio-temporal optimization. In ICCV, 2017.
[26]. Minwoo Park, Kyle Brocklehurst, Robert T Collins, and Yanxi Liu. Image de-fencing revisited in ACCV, 2010.
[27]. Abhijith Punnappurath and Michael S Brown. Reflection removal using a dual-pixel sensor. In CVPR, 2019.
[28]. Soumyadip Sengupta, Jinwei Gu, Kihwan Kim, Guilin Liu, David W Jacobs, and Jan Kautz. Neural inverse rendering of an indoor scene from a single image. In ICCV, 2019.
[29]. YiChang Shih, Dilip Krishnan, Fredo Durand, and William T Freeman. Reflection removal using ghosting cues. In CVPR, 2015.
[30]. Sudipta N Sinha, Johannes Kopf, Michael Goesele, Daniel Scharstein, and Richard Szeliski. Image-based rendering for scenes with reflections. ACM TOG, 31(4):100–1, 2012.
[31]. Deqing Sun, Xiaodong Yang, Ming-Yu Liu, and Jan Kautz. Pwc-net: Cnns for optical flow using pyramid, warping, and cost volume. In CVPR, 2018.
[32]. Yu Sun, Xiaolong Wang, Zhuang Liu, John Miller, Alexei A Efros, and Moritz Hardt. Test-time

training for out-of-distribution generalization. arXiv:1909.13231, 2019.

[33]. Richard Szeliski, Shai Avidan, and P Anandan. Layer extraction from multiple images containing reflections and transparency. In CVPR, 2000.

[34]. Renjie Wan, Boxin Shi, Ling-Yu Duan, Ah-Hwee Tan, and Alex C Kot. Benchmarking single-image reflection removal algorithms. In ICCV, 2017. 6

[35]. Zhou Wang, Alan C Bovik, Hamid R Sheikh, Eero P Simoncelli, et al. Image quality assessment: from error visibility to structural similarity. TIP, 13(4):600–612, 2004.

[36]. Kaixuan Wei, Jiaolong Yang, Ying Fu, David Wipf, and Hua Huang. Single image reflection removal exploiting misaligned training data and network enhancements. In CVPR, 2019.

[37]. Tianfan Xue, Baian Chen, Jiajun Wu, Donglai Wei, and William T Freeman. Video enhancement with the task-oriented flow. IJCV, 127(8):1106–1125, 2019.

[38]. Tianfan Xue, Michael Rubinstein, Ce Liu, and William T Freeman. A computational approach for obstruction-free photography. ACM TOG, 34(4):79, 2015.

[39]. Jie Yang, Dong Gong, Lingqiao Liu, and Qinfeng Shi. Seeing deeply and bidirectionally: A deep learning approach for single image reflection removal. In ECCV, 2018.

[40]. Renjiao Yi, Jue Wang, and Ping Tan. Automatic fence segmentation in videos of dynamic scenes. In CVPR, 2016.