



Mobile Image Restoration via Prior Quantization

Shiqi Chen^{a,2}, Jingwen Zhou^{a,2}, Menghao Li^{a,2}, Yueting Chen^{a,2}, Tingting Jiang^{b,**}

^aZhejiang University, Zheda Rd. 38, Hangzhou, 310000, Zhejiang, China

^bZhejiang Laboratory, Zhejiang Lab Nanhu Headquarters, Hangzhou, 311121, Zhejiang, China

Article history:

Image Restoration, Mobile ISP Systems, Computational Photography, Deep Learning

ABSTRACT

In the photograph of mobile terminal, image degradation is a multivariate problem, where the spectral of the scene, the lens imperfections, the sensor noise, and the field of view together contribute to the results. Besides eliminating it at the hardware level, the post-processing system, which utilizes various prior information, is significant for correction. However, due to the content differences among priors, the pipeline that directly aligns these factors shows limited efficiency and unoptimized restoration. Here, we propose a prior quantization model to correct the degradation introduced in the image formation pipeline. To integrate the multivariate messages, we encode various priors into a latent space and quantify them by the learnable codebooks. After quantization, the prior codes are fused with the image restoration branch to realize targeted optical degradation correction. Moreover, we propose a comprehensive synthetic flow to acquire data pairs in a relative low computational overhead. Comprehensive experiments demonstrate the flexibility of the proposed method and validate its potential to accomplish targeted restoration for mass-produced mobile terminals. Furthermore, our model promises to analyze the influence of various priors and the degradation of devices, which is helpful for joint soft-hardware design.

© 2023 Elsevier Ltd. All rights reserved.

1. Introduction

Any digital imaging system suffers from optical aberration and noise, and thus correcting these defects are necessary for accurate measurements [1]. Unfortunately, the optical degradation in image is affected by many factors in imaging pipeline, which is generally formulated as

$$J_e(h, w) = \int C_e(\lambda) \cdot [I_e(h, w, \lambda) * L_e(h, w, \lambda)] d\lambda + N_e(h, w), \quad (1)$$

here (h, w) indicates the coordinates of the pixel, λ is the wavelength, C_e and I_e denote the spectral response and the point spread function (PSF), L_e , J_e , and N_e are the latent sharp image, the observed image, and the measurement noise, respectively. Note that the subscript e indicates the measurement that represents the energy received by the sensor. Therefore, energy

dispersion and FoV clue of the optical system, spectral properties, and sensor noise together contribute to the deterioration expression on digital images. In other words, these factors serve as the priors to facilitate the correction.

Although there are many algorithms for image restoration, it still faces a few challenges for widespread application [2]. One issue is that the existing methods generally have a limited application scope, e.g., the deconvolution is inefficient in handling spatial-varying kernels, and the deep learning method is trained for a specific device according to the data [3]. Another challenge is the quality of restoration, which is generally unsatisfactory due to the lack of sufficient information [4]. An inherent defect of these unflexible methods is that they are incapable of mining the interaction between the digital pixel and other optical priors [5].

As mentioned above, the expression of optical degradation correlates with multiple factors [6]. Designing a general method to utilize these priors for targeted correction is an issue worth discussing. Meanwhile, this baseline can help analyze

**Corresponding author email: eagerjtt@zhejianglab.com

²Authors' emails: {chenshiqi, 22130055, 12130028, chenyt}@zju.edu.cn

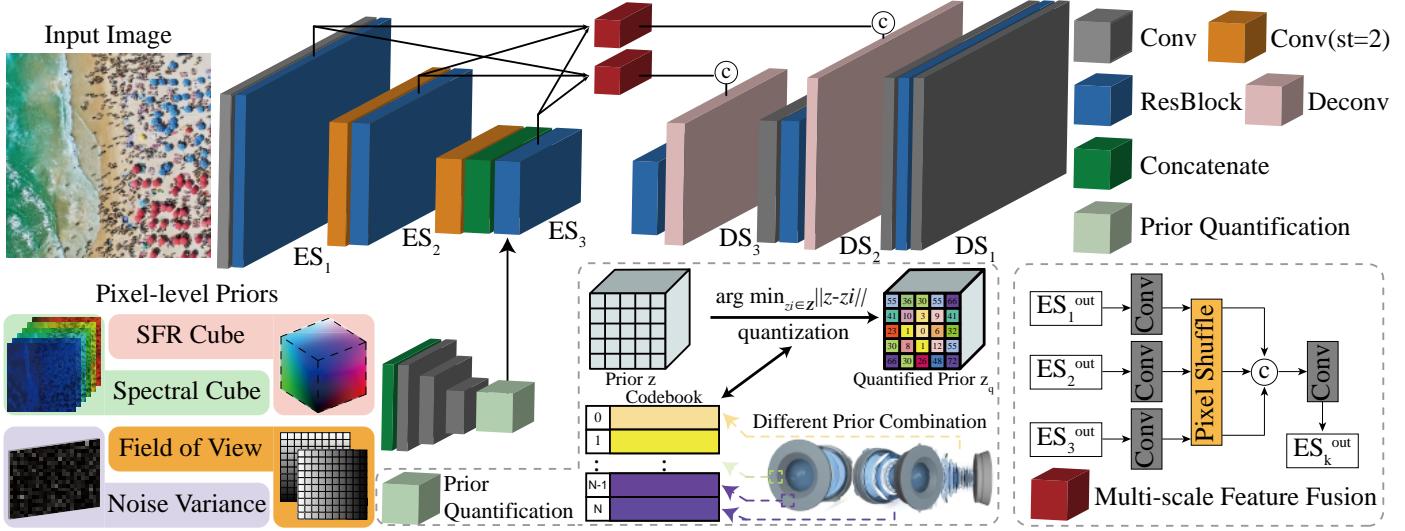


Fig. 1. The detail of the proposed prior quantization model. The layer configurations are illustrated with different colored blocks.

the correlation between different priors and image restoration, aiming to guide the co-design of the hardware configuration and the post-processing pipeline in the high-end imaging system.

In this letter, we propose a prior quantization model to correct the optical degradation influenced by multiple factors, where different priors corresponding to each factor are fed to the model. However, due to the content differences among the multimodal priors, they cannot directly integrate into the model for efficient post-processing. To this end, we encode the various priors into a high-dimensional latent space and characterize it by a learnable codebook. The learned code reduces the dimension of multimodal priors representation and models the global interrelations of auxiliary information. Then the model integrates the quantized prior representation into the image restoration branch by fusing it with features of different scales. Finally, we supervise the restoration in multiple scales to ensure that the cross-scale information used for fusion is accurate and valid. Instead of a black box, the model can adjust the input pixel-level prior according to the demands of the user, aiming to achieve a targeted development. Moreover, the learned codebook bridges the gap between different priors and restoration. Therefore, the proposed model has the potential to analyze the utilization of priors, which is meaningful for imaging system design. The contributions of this letter are summarized as follows:

- A prior quantization model is proposed to efficiently utilize various prior information and perform targeted restoration for mobile terminal.
 - We propose a comprehensive synthetic flow for generating the data pairs and their corresponding priors, significantly eliminating the overhead of data acquisition.

2. Method

Because the complexity of mobile degradation and the computational overhead limitation, the original intention of the proposed algorithm aims to effectively integrate the priors and

carry out a real-time restoration in the meanwhile. In the Sec. 2.1, we first illustrate the proposed prior quantization branch. Second, the image restoration branch is detailed in the Sec. 2.2. Besides the model of the network, the arrangement of data is significant for restoration, which is explained in the Sec. 2.3.

2.1. Prior Quantization Branch

As mentioned above, multiple factors contribute to the expression of optical degradation. Thus the model must possess the ability to perceive a complex combination and transform it into a simpler form. Directly engaging the image signal with individual priors is inefficient, where the model will spend substantial computing overhead on the auxiliary information. Therefore, we encode these priors into a high-dimension space and represent them with a learnable codebook. Since the aberration is relatively constant in the neighborhood, the optical prior of an image $x \in \mathcal{R}^{H \times W \times 3}$ can be represented by a spatial collection of codebook $z_q \in \mathcal{R}^{h \times w \times d_z}$, where d_z is the dimension of the code. To effectively learn such a codebook, we propose to exploit the spatial extraction capability of convolution and incorporate with the neural discrete representation learning [7]. First, we use CNNs to encode the priors into a latent space with the same spatial resolution ($h \times w$). Then, we embed the encoded features into the spatial code $\hat{z}_{ij} \in \mathcal{R}^{d_z}$ and quantify each code onto its closest entry in the discrete codebook $\mathcal{Z} = \{z_k\}_{k=1}^K \subset \mathcal{R}^{d_z}$:

$$z_q = \left(\arg \min_{z_k \in \mathcal{Z}} \| \hat{z}_{ij} - z_k \| \right) \in \mathcal{R}^{h \times w \times d_z}, \quad (2)$$

Third, we concatenate the quantized prior representation with the features in the restoration branch. And the subsequent Res-Block actively fused the image feature, where the prior after quantization provides clues of the spatial/channel importance to the reconstruction branch.

In most reconstruction models, the skip connections are only processed in one scale when the coarse-to-fine architecture is applied. Inspired by the dense connection between multi-scale

features, we implement a multi-scale feature fusion (MFF) module to integrate the quantified representation with the features from other scales [8]. This module receives the outputs of different encoding scales (ES_i) as inputs. After adjusting the channels of each ES_i^{out} by convolution, we use pixel shuffle to transform the encoded information into the same spatial resolution and then perform the fusion. The output of the MFF is delivered to its corresponding decoding scales (DS_i). More specifically, the MFFs of different scales are formulated as follows:

$$\begin{aligned} MFF_1 &= C_{DS1}(\{\mathcal{P}(C_{\times 1}(ES_1)), \mathcal{P}(C_{\times 2}(ES_2)), \mathcal{P}(C_{\times 4}(ES_3))\}), \\ MFF_2 &= C_{DS2}(\{\mathcal{P}(C_{\times 0.5}(ES_1)), \mathcal{P}(C_{\times 1}(ES_2)), \mathcal{P}(C_{\times 2}(ES_3))\}), \end{aligned} \quad (3)$$

where C is the operation of convolution, and the subscript refers to the ratio by which the number of input channels is to be multiplied after convolution. \mathcal{P} is the pixel shuffle (unshuffle) to rearrange the features of different scales into the same spatial resolution so that they be concatenated (denoted by $\{\cdot\}$ operation) together. In this way, each scale can perceive the encoded information of other scales, especially the lowest-scale features filtered by the quantified priors, resulting in improved restoration quality.

Here we discuss the loss function of our model. Due to the quantization operation in Eq. 2 is non-differentiable in back-propagation, the CNN encoder of priors cannot receive a gradient to optimize its parameters if the entire network is directly trained with end-to-end supervision. Fortunately, the strategy of the codebook alignment allows the gradient propagated from decoders to update these CNN encoders. Specifically, we keep the encoded priors approaching the vectors of the learnable codebook. In this way, the quantization progress like a gradient estimator, allowing the CNN encoder to estimate the backpropagation and update the parameters even when the gradient is truncated in training. Thus, we apply the codebook alignment loss to realize the optimization of encoders [9]:

$$\mathcal{L}_{align} = \|sg[\hat{z}_{ij}] - z_q\|_2^2 + \|\hat{z}_{ij} - sg[z_q]\|_2^2, \quad (4)$$

here $\hat{z}_{ij} = \{E_1(p_1), E_2(p_2), \dots, E_n(p_n)\}$, where p_i and E_i are the i^{th} prior and encoder, respectively. $\{\cdot\}$ is the operation to concatenate the encoded features. $sg[\cdot]$ denotes the stop-gradient operation to ensure the loss function only guide the encoded priors and the codebook vectors approaching.

2.2. Image Restoration Branch

Because our network relies on the refinement in a coarse-to-fine manner, we supervise the image reconstruction on different scales. For the supervision of image content, we find that the L1 loss performs better on quantitative metrics for image degradation correction. However, since the content loss only measures the pixel-level difference, the model does a good job restoring the low-frequency information (such as brightness and color, etc.) while performing poorly in restoring the textures of the scene. To prevent the limitation of pixel-level supervision, we adopt the supervision in the Fourier domain as an auxiliary loss [8]. Compared with the gradient constraints (e.g., total variation) or the feature similarity on the pre-trained model (e.g.,

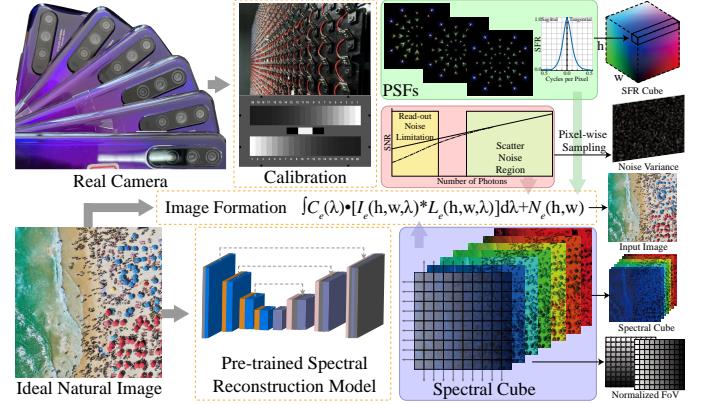


Fig. 2. The synthetic flow of the training data and the priors.

perceptual loss), this item provides more information on different spatial frequencies. The loss $\mathcal{L}_{content}$ is formulated as follows:

$$\mathcal{L}_{content} = \mathcal{L}_{rec} + \lambda \mathcal{L}_{freq} = \sum_{k=1}^K \|\hat{I}_k - I_k\|_1 + \lambda \|\mathcal{F}(\hat{I}_k) - \mathcal{F}(I_k)\|_1, \quad (5)$$

here the output of the k -th scale is I_k and its corresponding ground-truth is \hat{I}_k , where K is the number of scales. The content loss in each scale is the average on the total elements, whose number is denoted as t_k . where \mathcal{F} is the fast Fourier transform (FFT) operation that performs on the predicted images of different scales. The overall loss function of the proposed model is determined as follows:

$$\mathcal{L}_{total} = \mathcal{L}_{rec} + \mathcal{L}_{align} + \lambda \mathcal{L}_{freq}. \quad (6)$$

where the λ is experimentally set to 0.1.

2.3. Synthetic Flow for image pairs and priors

Since the image degradation in mobile terminal is highly correlated to multiple factors, we construct a comprehensive dataset for correction, where every scenes and their corresponding priors are obtained. The detailed synthetic flow of the training data and its priors is shown in Fig. 2. First, we calibrate the PSFs and the noise parameters of the mobile cameras. The PSFs are used to simulate the degradation (as I_e) and calculate the SFR cube. Using the parameters calibrated by gray-scale board, including read-out and scatter noise factors, we determine the noise distribution (as N_e) in image formulation and the prior of noise. Second, we use the pre-trained MST++ [10] to convert RGB images to hyperspectral cube. This data (refer as L_e) will be integrated according to sensor response C_e for later imaging simulation. Moreover, the spectral cube is regarded as the spectral prior for later quantization. Third, the FoV prior is acquired by normalizing the (h, w) pixel coordinates to $[-1, 1]$ [11, 12]. Finally, these data is engaged with the degradation of image formation pipeline, where the procedure is the same as Eq. 1. In implementation, we use six samples of Honor 20 and two terminals of iPhone 12 for data preparation, where the data-pairs of three Honor 20 and one iPhone 12 make up the training dataset, the rests consist of the test dataset.

Table 1. Performance of the proposed model and other competing methods on synthetic and real data. The percentage denotes the relative improvement compared with the best models (PSNR first turn to RMSE, other metrics have no change).

Method	TestSet	Synthetic Evaluation				Real Evaluation		
		PSNR↑	SSIM↑	VIF↑	LPIPS↓	BRISQUE↓	NIQE↓	
SRN		29.79 (56.6%)	0.9247 (29.1%)	0.8221 (11.3%)	3.428 (48.2%)	54.49 (16.1%)	5.686 (21.3%)	
IRCNN		30.58 (48.0%)	0.9289 (24.5%)	0.8314 (10.0%)	3.151 (43.7%)	53.47 (14.5%)	5.572 (19.7%)	
Self-Deblur		32.23 (24.0%)	0.9353 (17.5%)	0.8544 (7.05%)	2.766 (35.8%)	50.28 (9.03%)	5.323 (16.0%)	
GLRA		32.47 (19.6%)	0.9347 (18.2%)	0.8662 (5.60%)	2.457 (27.8%)	49.74 (8.04%)	5.476 (18.3%)	
FDN		33.27 (3.39%)	0.9402 (12.2%)	0.8943 (2.28%)	1.944 (8.69%)	47.42 (3.54%)	5.038 (11.2%)	
Deep Wiener		32.02 (27.6%)	0.9463 (5.7%)	0.9007 (1.55%)	1.928 (7.94%)	46.52 (16.8%)	4.943 (9.44%)	
Ours		33.42 (0.0%)	0.9517 (0.0%)	0.9147 (0.0%)	1.775 (0.0%)	45.74 (0.0%)	4.476 (0.0%)	

3. Experimental Results

We design the extensive experiments for evaluating the proposed method. In Sec. 3.1, the details of experiment are first illustrated. Second, we compare the proposed model with other state-of-the-art (SOTA) method in Sec. 3.2. Third, we carry out an ablation study in Sec. 3.3. At last, the real restoration and the correlation analyse are presented in Sec. 3.4 and Sec. 3.5.

3.1. Datasets, Metrics, and Training Settings

As illustrated above, we use the natural images in DIV2K, DIV8K, and FiveK to synthetic the training data-pairs and priors [13, 14, 15]. The training dataset consists of 800 image pairs. And the resolution of image is rescaled to 3000×4000 to align with the real-world evaluation. In the case of real-world evaluation, we capture many photographs with multiple mobile terminal of Huawei Honor 20 and iPhone 12. As for the metrics, the PSNR, SSIM, VIF [16], and LPIPS [17] evaluate the model with reference. The BRISQUE, NIQE [18] are used to assess the restoration of the captrued photographs. In the training, we crop the whole image to 256×256 pixels and form minibatches of 8. For a fair comparisons, the priors are concatnated with the images and fed into the competing models. We opti-mize the model with Adam in $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The initial learning rate is 10^{-4} and then halved every 2×10^5 iter-ations, and the optimization runs 50 epoch on the whole. The implement platform, the specific configuration of the proposed model are provided in the **supplementary material**.

3.2. Quantitative Assessment to SOTA Methods

We compare the proposed prior quantization model with the competing algorithms designed for image restoration. All these methods are retrained with the same dataset until convergence. In the assessment of spatial-various aberration correction, we evaluate the quantitive indicators on the synthetic dataset. Tab. 1 reports the average performance of various approaches on the testsets (the standard variances of evaluations are listed in **supplementary material**). The blind methods (SRN [19], IRCNN [20], SelfDeblur [21], GLRA [22]) successfully deal with the optical degradation of one camera. But there are various mobile cameras with different optical degradation, and the blind manner fails to acquire the precise degradation clue only from the feature of the image. The non-blind approaches (FDN [23], Deep Wiener [24]), which feed the pre-trained models with PSFs to adapt to a specific camera, have a similar idea to ours. However, since the PSFs is a high-dimensional representation



Fig. 3. Visual ablation study on the Prior Quantization.

Table 2. Quantitative ablation study for critical modules on synthetic (PSNR↑ / SSIM↑) and real (BRISQUE↓ / NIQE↓) data.

Evaluations	w/o PQ	w/o MSO	Ours
Synthetic data	32.29 / 0.9353	33.34 / 0.9446	33.42 / 0.9517
Real data	48.01 / 5.295	45.93 / 4.922	45.74 / 4.476

of degradation, only a few representative PSFs are selected to be fed into the model (only 5 PSFs in [24]), where the spatial relationship of the PSFs across the whole FoV is abandoned. Otherwise, the complexity will be extremely high. Different from this method, we use the rearranged SFR cube to represent the dispersion for each FoV, which is a lower-dimensional representation of degradation. Therefore, with relatively lower computational overhead, our model can obtain pixel-level guidance to achieve better restoration. For the detailed comparisons on different datasets, we refer readers to the **supplementary**.

3.3. Ablation Study

prior quantization (denoted as PQ): We retrain the model without prior quantization branch to verify its effectiveness. For a fair comparisons, the corresponding priors are concatenate with the image feature in the scale with same spatial resolution. The comparison is shown in the Fig. 3, we note that the restoration without quantization will cause a slight color cast, which means that directly integrating various priors with the image feature is inefficient, even counterproductive. Besides the visual comparison, the enhancements of indicators in Table 2 show that prior quantization is significant for the restoration.

multi-scale fusion and supervision (denoted as MSO): Because whether deploy the multi-scale operations show little difference on visualization, we mainly evaluate their validity by quantitative comparisons. As shown in Table 2, the performance of the proposed model is slightly decreased without the fusion and supervision in different scales. Thus, cross-scale information is useful for the image restoration in mobile terminal.

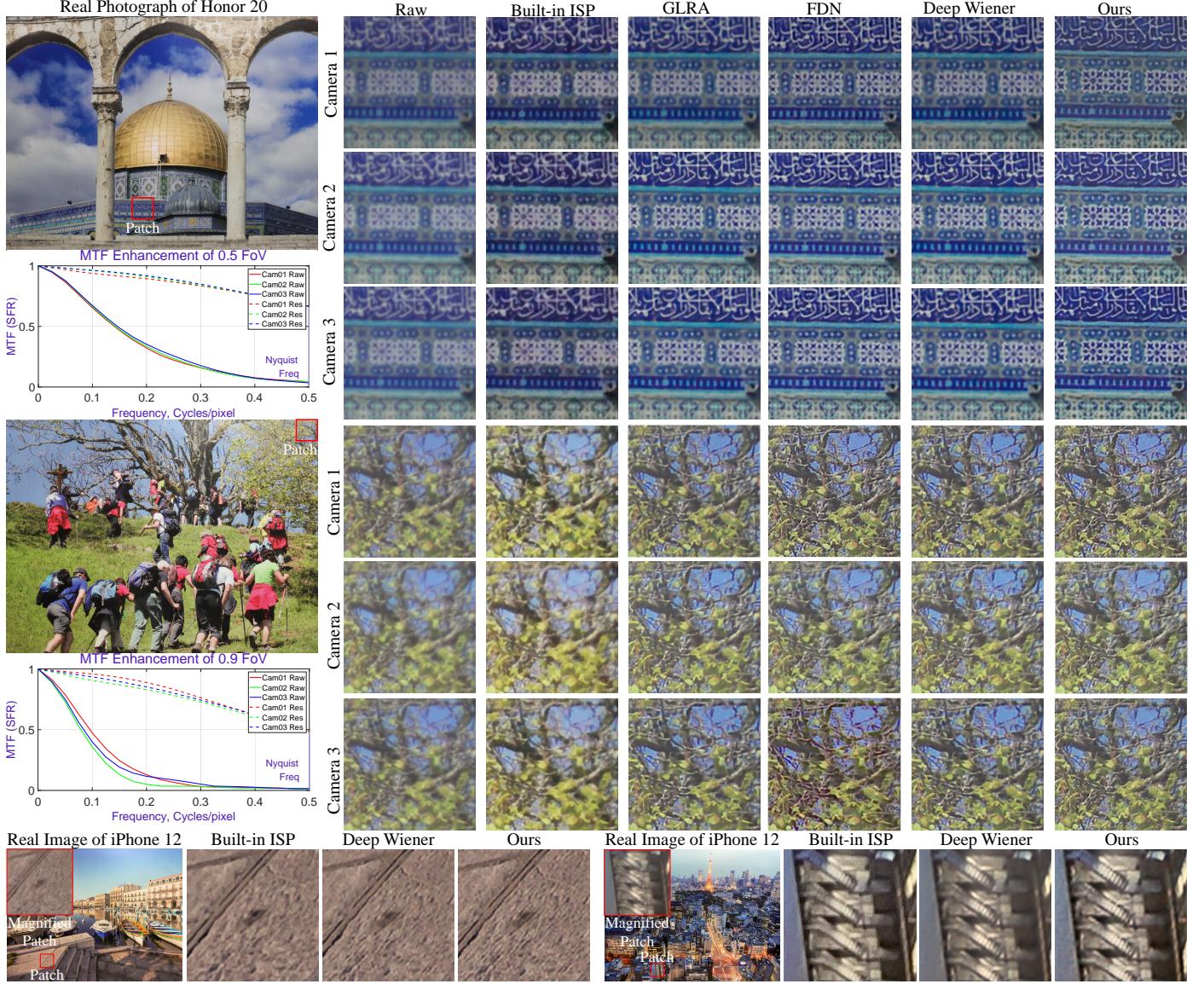


Fig. 4. Real image restoration comparison, where the magnified position is highlighted in red. We show multiple samples of Honor 20 and iPhone 12 to evaluate the generalization of the proposed method. The MTF enhancements of different FoVs are plotted.

3.4. Real Restoration Comparisons

In this section, we test these algorithms with the degraded photographs taken from real cameras, and the results are visualized in Fig. 4. Since the priors acquired by the Deep Wiener do not correspond to the pixel neighborhood, the model needs attention mechanism to determine the specific degradation of the input. These overheads increase the processing burden, resulting in the failure to obtain efficient restoration. Compared with other algorithms and the built-in ISP, our model efficiently integrates various priors and characterizes them with the quantized codebook vectors. Therefore, our method reduces the complexity of multi-task post-processing, and achieves a comprehensive improvement in image quality. Another advantage of our non-blind model is that it can be designed as a post-processing system with better generalization. In our experiment, we use the model trained on the simulation dataset to post-process the real images taken by various mobile cameras. The results in the left

and right of Fig. 4 demonstrate that the model pre-trained on synthetic data achieves perfect generalization ability on specific devices (Honor 20 and iPhone 12), where fine-tuning is not necessary. Therefore, the plug-and-play feature indicates that the learned model is promising to replace the camera-specific ISP system with a flexible model. For the detailed comparisons, we refer the readers to the **supplementary material** for more restorations on different manufactured samples and more the quantitative results of system imaging quality.

3.5. Prior Correlation Analyse

We separately evaluate **the influence of the pixel-level prior on the restored image quality**, especially the cases when priors are mismatched. In specifically, the priors are shuffled (detailed in the **supplementary**) then fed into the learned model in inference. The representative results are collated in Fig. 5. To better highlight the role played by each prior, we shown the

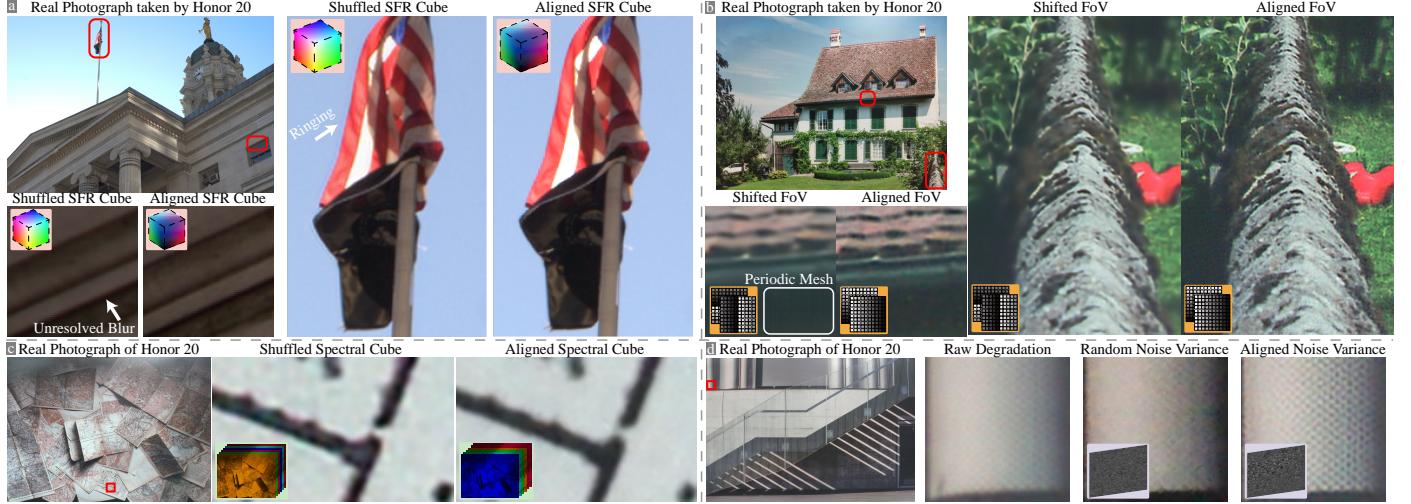


Fig. 5. Impact of each pixel-level prior on the image restoration. We show the influences on restored images that are carried out by the mismatch of pixel-level priors, where (a) is the SFR Cube, (b) is the FoV, (c) is the Spectral Cube, and (d) is the Noise Variance.

zoomed patches from the restored real photographs and the corresponding prior information. In the following, we will analyze the influence of each prior separately,

- **SFR Cube:** here the SFR sequence in the spatial dimension will exchange position randomly, which means that the SFR that indicates severe degradation in the edge of FoV has the possibility to be assigned to the image center. This will cause overcorrection (accompany with pseudo texture). And when the degradations represented by SFRs does not correspond, the images will leave unresolved blur in the edge. In Fig. 5 (a), we find that the extremely low SFR information can cause the ghosting of flags near the image center, while the blur on ledges remains unresolved because of the SFR mismatch.

- **FoV:** similar to SFR Cube, the image restoration with the FoV prior exhibits highly spatial correlation. As shown in Fig. 5 (b), the restoration of the swapped FoV is unsatisfactory in the border of image, where the details of the trunk and the leaf are blurred. While in the center of the photograph, the shuffled FoV will introduce a periodic grid in the areas with weak texture (under the roof).

- **Spectral Cube:** for the optical imaging system in visible light band, the chromatic aberration is highly correlated with the spectrum of scene. However, the learned model is difficult to predict the accurate degradation in RGB when the spectral prior is inaccurate. Therefore, as shown in Fig. 5 (c), unresolved color fringing will appear when feed the shuffled spectral cube, especially in high-contrast region.

- **Noise Variance:** the influence of the noise variance is intuitive. As shown in Fig. 5 (d), the randomly sampled noise variance can't provide the targeted benefits for denoising, while the signal-related variance performs well in the restoration of real captured photographs.

These experiments demonstrate that each prior plays key roles in the mobile image restoration. From the other per-

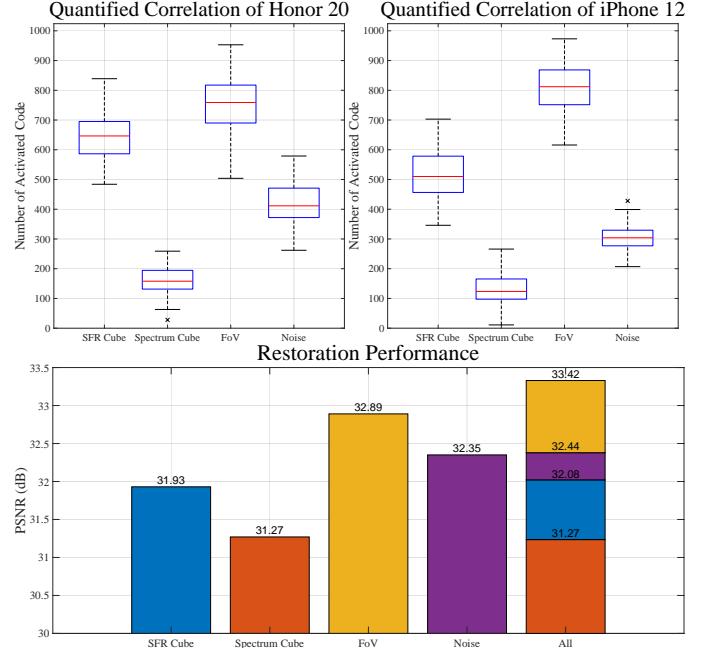


Fig. 6. The correlation between priors and image degradation correction in different mobile cameras.

spective, these pixel-level priors can be regarded as the vision prompts to guide the learned model realize targeted correction. In the **supplementary**, our method are compared with the prompt-based restoration on vision transformer, where we demonstrate that the prior quantization is more efficient when the amount of parameters is limited.

Unfortunately, obtaining so much assistance in real-time imaging comes at an unaffordable expense. We use the learned model to **analyze the proceeds of introducing each prior into the restoration**. When evaluating a specific prior, we zero out all other auxiliaries except this one and test the learned model with the raw photographs, then count the number of the activated codes in the inference process. The larger number of

activated codebooks indicates the more relevance between this prior and the restoration. The result of the assessment is shown at the top of Fig. 6. We find that the SFR and FoV priors will activate more codes in inference, which means they are highly correlated with aberration and play critical roles in correction. Note that the restoration of the Honor 20 activates more code in the SFR and noise priors evaluation when compared with the iPhone 12, which attributes to its uneven optical aberration of the optics and lower SNR of the sensor. Therefore, the learned model could hint the utilization of different priors, providing a brand new topic for camera assessment. Moreover, we evaluate the correlation between the restoration indicators and the priors in the bottom of Fig. 6. Different from the number of activated codes, the noise is more substantial on the restoration performance, which may put down to the calculation of PSNR that assesses the absolute error from pixel scale.

4. Conclusion

In summary, we develop a deep learning model and its accompanying data generation pipeline, which utilize multiple priors for optical degradation correction. Even though the content of the priors varies a lot from each other, the proposed quantization strategy efficiently fuses these factors and guides the correction of image degradation. Comprehensive experiments show that integrating all the related information does benefit the restoration of optical degradation, and the proposed model can generalize to different mobile devices without fine-tuning. Moreover, the learned model quantifies the correlation and significance of different priors for the correction procedure. Thus, the aberration correction in the post-processing system can be more efficient when the critical clues are obtained. In the future, we will put efforts into engaging the proposed model with different imaging devices, aiming to realize efficient correction and better generalization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was sponsored by the Civil Aerospace Pre-Research Project D040107, the National Natural Science Foundation of China (501100001809) 61424080211, and the Key Research Project of Zhejiang Lab 2021MH0AC01.

References

- [1] T. Yue, J. Suo, J. Wang, X. Cao, Q. Dai, Blind optical aberration correction by exploring geometric and visual priors, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2015) 1684–1692.
- [2] Z. Zhang, X. Song, X. Sun, V. Stojanovic, Hybrid-driven-based fuzzy secure filtering for nonlinear parabolic partial differential equation systems with cyber attacks, International Journal of Adaptive Control and Signal Processing 37 (2) (2023) 380–398.
- [3] T. Eboli, J.-M. Morel, G. Facciolo, Fast two-step blind optical aberration correction, European Conference on Computer Vision (2022) 693–708.
- [4] L. Shen, H. Tao, Y. Ni, Y. Wang, V. Stojanovic, Improved yolov3 model with feature map cropping for multi-scale road object detection, Measurement Science and Technology 34 (4) (2023) 045406.
- [5] S. Chen, T. Lin, H. Feng, Z. Xu, Q. Li, Y. Chen, Computational optics for mobile terminals in mass production, IEEE Transactions on Pattern Analysis and Machine Intelligence (2022) 1–16doi:10.1109/TPAMI.2022.3200725.
- [6] X. Song, N. Wu, S. Song, V. Stojanovic, Switching-like event-triggered state estimation for reaction-diffusion neural networks against dos attacks, Neural Processing Letters (2023) 1–22.
- [7] A. Van Den Oord, O. Vinyals, et al., Neural discrete representation learning, Advances in neural information processing systems 30 (2017).
- [8] S.-J. Cho, S.-W. Ji, J.-P. Hong, S.-W. Jung, S.-J. Ko, Rethinking coarse-to-fine approach in single image deblurring, Proceedings of the IEEE/CVF international conference on computer vision (2021) 4641–4650.
- [9] J. Duan, L. Chen, S. Tran, J. Yang, Y. Xu, B. Zeng, T. Chilimbi, Multi-modal alignment using representation codebook, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2022) 15651–15660.
- [10] Y. Cai, J. Lin, Z. Lin, H. Wang, Y. Zhang, H. Pfister, R. Timofte, L. Van Gool, Mst++: Multi-stage spectral-wise transformer for efficient spectral reconstruction, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2022) 745–755.
- [11] S. Chen, H. Feng, K. Gao, Z. Xu, Y. Chen, Extreme-quality computational imaging via degradation framework, Proceedings of the IEEE/CVF International Conference on Computer Vision (2021) 2632–2641.
- [12] Q. Sun, C. Wang, Q. Fu, X. Dun, W. Heidrich, End-to-end complex lens design with differentiable ray tracing, ACM Trans. Graph. 40 (4) (jul 2021). doi:10.1145/3450626.3459674
URL <https://doi.org/10.1145/3450626.3459674>
- [13] E. Agustsson, R. Timofte, Ntire 2017 challenge on single image super-resolution: Dataset and study, in: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2017.
- [14] S. Gu, A. Lugmayr, M. Danelljan, M. Fritzsche, J. Lamour, R. Timofte, Div8k: Diverse 8k resolution image dataset, in: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), IEEE, 2019, pp. 3512–3516.
- [15] V. Bychkovsky, S. Paris, E. Chan, F. Durand, Learning photographic global tonal adjustment with a database of input / output image pairs, in: The Twenty-Fourth IEEE Conference on Computer Vision and Pattern Recognition, 2011.
- [16] H. R. Sheikh, A. C. Bovik, Image information and visual quality, IEEE Transactions on image processing 15 (2) (2006) 430–444.
- [17] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, O. Wang, The unreasonable effectiveness of deep features as a perceptual metric, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 586–595.
- [18] A. Mittal, R. Soundararajan, A. C. Bovik, Making a “completely blind” image quality analyzer, IEEE Signal processing letters 20 (3) (2012) 209–212.
- [19] X. Tao, H. Gao, X. Shen, J. Wang, J. Jia, Scale-recurrent network for deep image deblurring, Proceedings of the IEEE conference on computer vision and pattern recognition (2018) 8174–8182.
- [20] K. Zhang, W. Zuo, S. Gu, L. Zhang, Learning deep cnn denoiser prior for image restoration, Proceedings of the IEEE conference on computer vision and pattern recognition (2017) 3929–3938.
- [21] D. Ren, K. Zhang, Q. Wang, Q. Hu, W. Zuo, Neural blind deconvolution using deep priors, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2020) 3341–3350.
- [22] W. Ren, J. Zhang, L. Ma, J. Pan, X. Cao, W. Zuo, W. Liu, M.-H. Yang, Deep non-blind deconvolution via generalized low-rank approximation, Advances in neural information processing systems 31 (2018).
- [23] J. Kruse, C. Rother, U. Schmidt, Learning to push the limits of efficient fft-based image deconvolution, Proceedings of the IEEE International Conference on Computer Vision (2017) 4586–4594.
- [24] T. Lin, S. Chen, H. Feng, Z. Xu, Q. Li, Y. Chen, Non-blind optical degradation correction via frequency self-adaptive and finetune tactics, Opt. Express 30 (13) (2022) 23485–23498. doi:10.1364/OE.458530.
URL <https://opg.optica.org/oe/abstract.cfm?URI=oe-30-13-23485>