[Supplementary Material] QueryCDR: Query-based Controllable Distortion Rectification Network for Fisheye Images

Pengbo Guo[1,2], Chengxu Liu[2,3], Xingsong Hou[2(✉)], and Xueming Qian[2,3]

[1] School of Software Engineering, Xi'an Jiaotong University, Xi'an, China
[2] Xi'an Jiaotong University, Xi'an, China
[3] Shaanxi Yulan Jiuzhou Intelligent Optoelectronic Tech. Co., Ltd, Xi'an, China
{guopengbo,chengxuliu}@stu.xjtu.edu.cn, {houxs,qianxm}@mail.xjtu.edu.cn

In this supplementary material, Sec. 1 provides the details of the synthetic fisheye image datasets. Sec. 2 conducts further ablation studies on learnable query and training strategy. Finally, Sec. 3 shows more quantitative and qualitative comparison results.

1 Fisheye Model and Data Synthesis

Due to the extreme difficulty in obtaining a large number of fisheye images and their corresponding labels in the real world, using distortion models to synthesize fisheye images has become the mainstream choice [2, 5, 6, 9]. Therefore, we follow existing methods to use the most commonly used polynomial model [3] to generate distorted fisheye images, described as follows,

$$\theta_u = \sum_{i=1}^{n} k_i \theta_d^{2i-1}, \quad n = 1, 2, 3, 4, \dots, \tag{1}$$

where θ_u represents the angle of incident light, θ_d is the angle of the light after passing through the lens, k_i is the i-th distortion parameter, and n is the number of distortion parameters. According to the pinhole camera projection model, we can get $r_d = f\theta_d$, where f is the camera's focal length, and r_d is the $L2$ distance between the image center and any point in the original image. Further simplifying with $r_u = f \tan\theta_u$, we obtain $\theta_u = \arctan \frac{r_u}{f} \approx \frac{r_u}{f}$. Therefore, the relationship between r_u and r_d in the polynomial model is as follows,

$$r_u = f \sum_{i=1}^{n} k_i r_d^{2i-1}. \quad n = 1, 2, 3, 4, \dots \tag{2}$$

Furthermore, we integrate f into k_i to obtain the final polynomial model,

$$r_u = \sum_{i=1}^{n} k_i r_d^{2i-1}. \quad n = 1, 2, 3, 4, \dots \tag{3}$$

Table 1: Ablation study of the transferability of the learnable query.

Method	PSNR									
	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	Avg
PCN	14.93	17.43	18.43	18.86	18.86	**18.88**	**18.74**	17.35	18.26	17.97
+Learnable Query	**15.21**	**17.94**	**18.93**	**19.26**	**19.20**	18.87	18.62	**17.53**	**19.28**	**18.32**

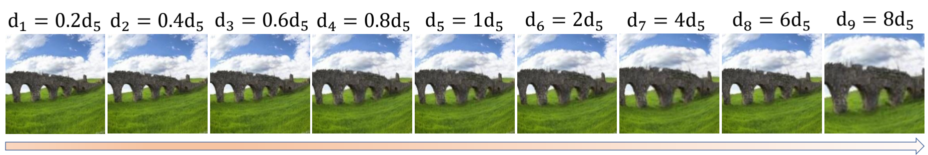

Fig. 1: Given different distortion parameters, the polynomial model synthesizes images with different degrees of distortion.

To synthesize fisheye image datasets with varying distortion degrees, we define the parameter setting used in previous methods [2, 5, 8, 9] as the general distortion degree d_5. Subsequently, we amplify and reduce d_5 by four degrees to get different degrees of distortion, respectively. To be specific, in order to set reasonable and distinguishable distortion degrees, we multiply the d_5's distortion parameters by 0.2, 0.4, 0.6, 0.8, 1, 2, 4, 6, and 8, getting the corresponding distortion degrees $\{d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8, d_9\}$. Then the polynomial model is employed using different parameters to synthesize images with different degrees of distortion, as shown in Fig. 1.

2 More Ablation Studies

2.1 Ablation Study of Learnable Query's Transferability

To further explore the generalization ability of our learnable query, and whether it can transfer learned rectification knowledge to guide other uncontrollable methods [2,5,6,8,9] for controllable rectification. We directly applied the trained query set from QueryCDR to multiply each layer's features of the uncontrollable PCN [8]. As shown in Tab. 1, we are delighted to find that the learnable query enables the uncontrollable rectification network to have the generalization ability to deal with distortions to some extent, manifested by performance improvements across multiple distortion degrees. This further proves the effective controlling ability of our learnable query.

2.2 Ablation Study of Different Training Strategy

To validate the effectiveness of our proposed two-stage training strategy (in Sec. 3.4 of the main paper), we compared the performance of training our

Table 2: Performance comparison of different training strategies.

Method	PSNR									
	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	Avg
Only Pre-training	19.52	20.25	19.97	19.43	19.89	20.07	20.03	18.76	20.15	19.78
Only Fine-tuning	19.96	20.15	20.30	20.22	20.46	20.51	20.41	18.80	20.15	20.11
Two-Stage	**20.01**	**20.29**	**20.39**	**20.41**	**20.72**	**20.81**	**20.58**	**19.11**	**20.53**	**20.32**

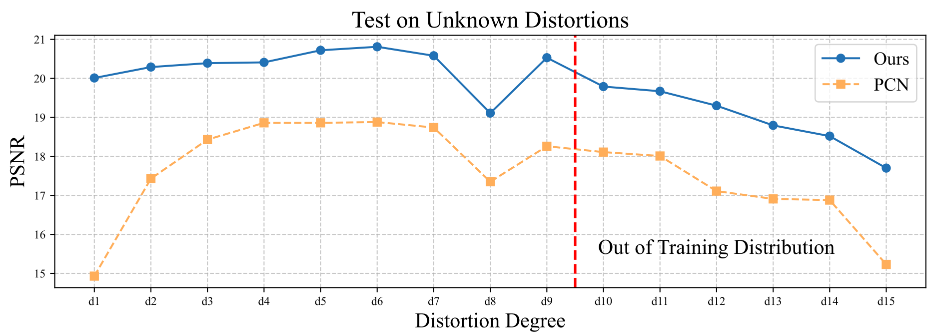

Fig. 2: Performance evaluation on out-of-distribution distortions.

QueryCDR solely with coarse-grained distortion pre-training and solely with fine-grained distortion fine-tuning. As shown in Tab. 2, compared to single-stage training methods, our two-stage training strategy achieves higher performance while converging faster. Such design also alleviates the requirement of controllable rectification models for data with multiple distortion degrees. By pre-training on a large amount of easy-to-obtain single distortion degree data, and then fine-tuning on a small amount of hard-to-obtain data with multiple distortion degrees, we successfully address the problem of degraded performance due to insufficient data in practical applications.

2.3 Ablation Study of Out-of-Distribution Distortion

To evaluate the robustness of our method on unknown distortions, *i.e.*, the degrees out of training distribution, we further explore the generalization ability on images with severer distortions. As shown in Fig. 2, QueryCDR still maintains higher performance compared to other rectification method [8].

3 More Results

3.1 More Qualitative Results

To further verify the effectiveness of our QueryCDR, we conducted experiments on fisheye image datasets based on the original images from Places2 [10] dataset, as shown in Tab. 3. Our QueryCDR demonstrates superior performance across all degrees of distortion on the Places2 dataset [10], further confirming its controllable rectification capability across different scenes and distortion degrees.

Table 3: Quantitative comparison (PSNR (dB)↑, SSIM ↑) on Places2 [10] fisheye image dataset with varying distortion degrees. Red indicates the best and blue indicates the second best performance (best viewed in color).

Method	PSNR									
	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	Avg
SC [7]	10.85	10.60	10.45	10.40	10.35	10.22	10.07	10.11	10.10	10.35
DeepCalib [1]	13.43	12.01	11.37	11.03	10.83	10.23	9.97	9.91	10.02	10.98
Blind [4]	10.83	10.83	10.50	10.45	10.65	10.31	10.40	10.19	10.24	10.49
DR-GAN [5]	16.89	18.35	19.03	19.38	19.27	20.15	20.37	20.24	19.90	19.28
PCN [8]	15.13	17.02	18.33	18.89	19.45	20.81	21.01	20.61	20.15	19.04
QueryCDR	20.89	21.26	21.14	21.46	21.07	21.62	22.02	21.52	21.31	21.36
Method	SSIM									
	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	Avg
SC [7]	0.153	0.144	0.136	0.135	0.133	0.128	0.123	0.126	0.128	0.134
DeepCalib [1]	0.375	0.298	0.267	0.251	0.242	0.219	0.212	0.211	0.221	0.255
Blind [4]	0.246	0.244	0.237	0.239	0.231	0.230	0.243	0.249	0.259	0.242
DR-GAN [5]	0.411	0.453	0.469	0.475	0.468	0.485	0.489	0.483	0.476	0.468
PCN [8]	0.427	0.521	0.589	0.605	0.630	0.693	0.705	0.686	0.671	0.575
QueryCDR	0.678	0.698	0.701	0.713	0.704	0.723	0.740	0.720	0.716	0.710

3.2 More Visual Results

We show more visual results among the proposed QueryCDR and other rectification methods [1, 4, 5, 7, 8] in Fig. 3 and Fig. 4. Moreover, we provide further comparisons with the second-best approach [8] to demonstrate the details in the rectified images, as shown in Fig. 5 and Fig. 6.

References

1. Bogdan, O., Eckstein, V., Rameau, F., Bazin, J.C.: Deepcalib: A deep learning approach for automatic intrinsic calibration of wide field-of-view cameras. In: Proceedings of the 15th ACM SIGGRAPH European Conference on Visual Media Production. pp. 1–10 (2018)
2. Feng, H., Wang, W., Deng, J., Zhou, W., Li, L., Li, H.: Simfir: A simple framework for fisheye image rectification with self-supervised representation learning. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 12418–12427 (2023)

3. Kannala, J., Brandt, S.S.: A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. IEEE transactions on pattern analysis and machine intelligence **28**(8), 1335–1340 (2006)
4. Li, X., Zhang, B., Sander, P.V., Liao, J.: Blind geometric distortion correction on images through deep learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4855–4864 (2019)
5. Liao, K., Lin, C., Zhao, Y., Gabbouj, M.: Dr-gan: Automatic radial distortion rectification using conditional gan in real-time. IEEE Transactions on Circuits and Systems for Video Technology **30**(3), 725–733 (2019)
6. Liao, K., Lin, C., Zhao, Y., Xu, M.: Model-free distortion rectification framework bridged by distortion distribution map. IEEE Transactions on Image Processing **29**, 3707–3718 (2020)
7. Santana-Cedrés, D., Gomez, L., Alemán-Flores, M., Salgado, A., Esclarín, J., Mazorra, L., Alvarez, L.: An iterative optimization algorithm for lens distortion correction using two-parameter models. Image Processing On Line **6**, 326–364 (2016)
8. Yang, S., Lin, C., Liao, K., Zhang, C., Zhao, Y.: Progressively complementary network for fisheye image rectification using appearance flow. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 6348–6357 (2021)
9. Yang, S., Lin, C., Liao, K., Zhao, Y.: Dual diffusion architecture for fisheye image rectification: Synthetic-to-real generalization. arXiv preprint arXiv:2301.11785 (2023)
10. Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., Torralba, A.: Places: A 10 million image database for scene recognition. IEEE transactions on pattern analysis and machine intelligence **40**(6), 1452–1464 (2017)

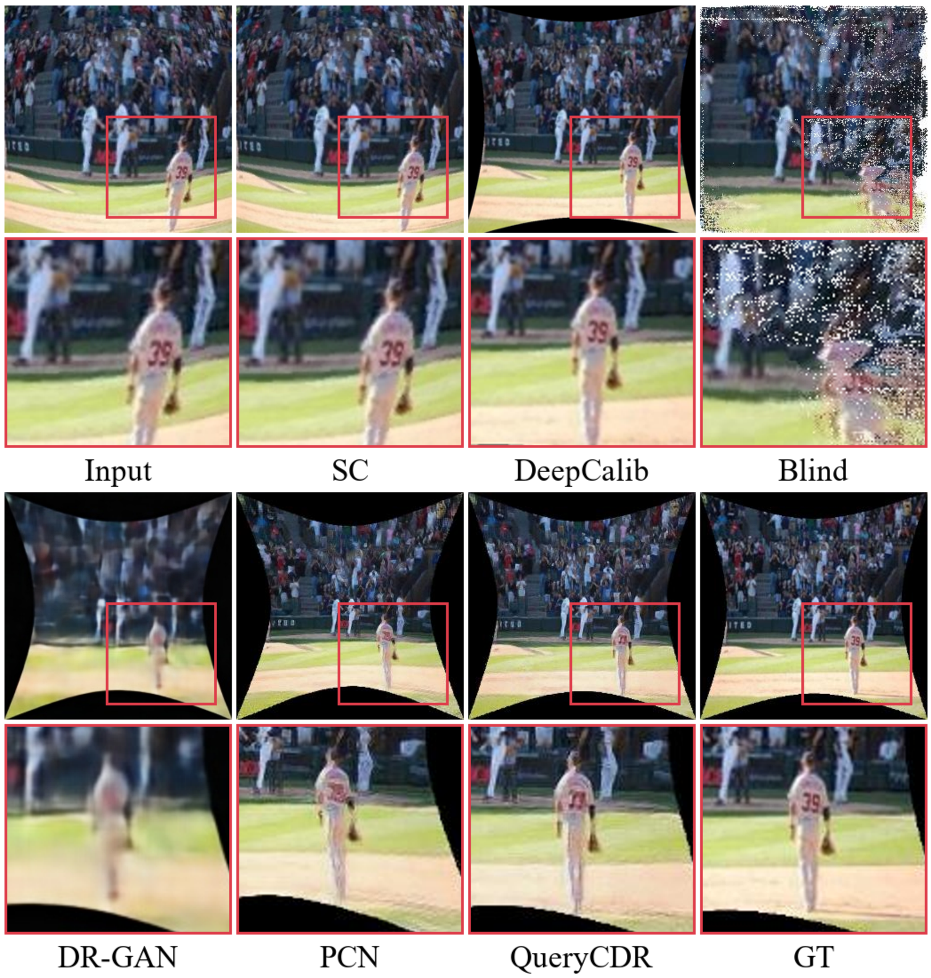

Fig. 3: The qualitative comparison with other rectification methods.

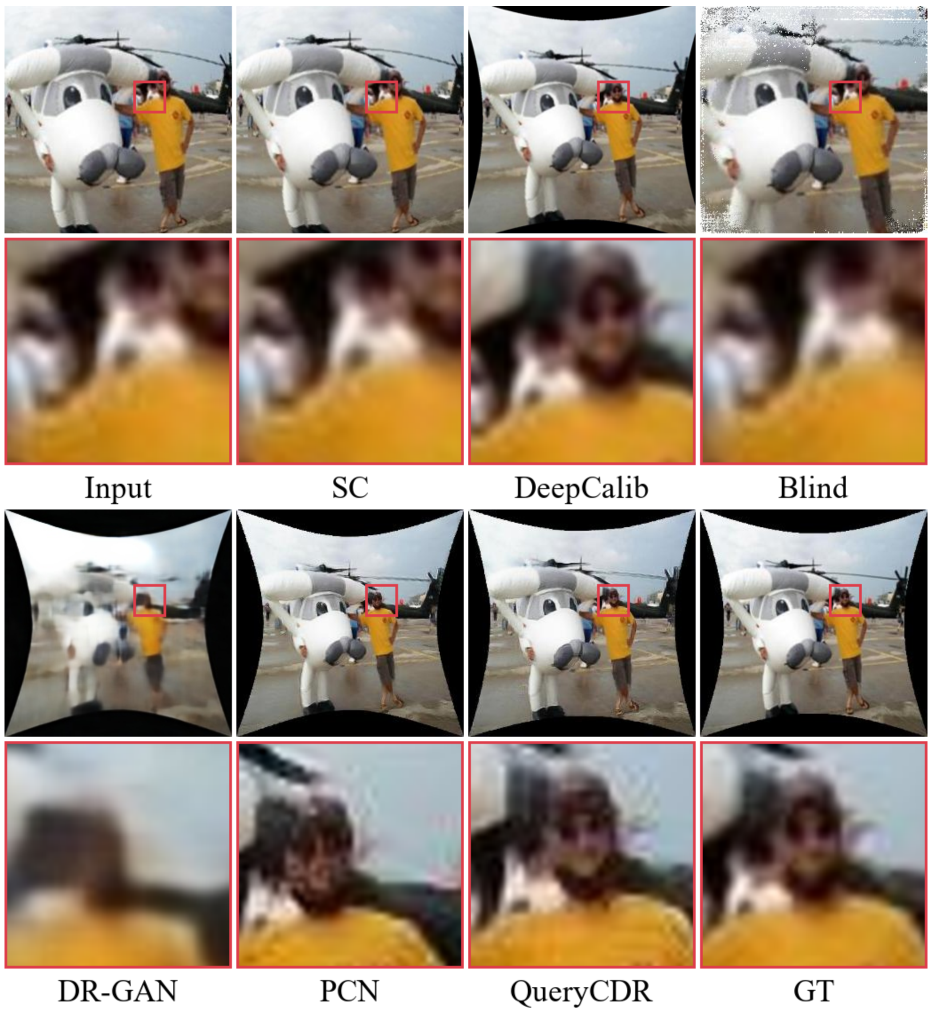

Fig. 4: The qualitative comparison with other rectification methods.

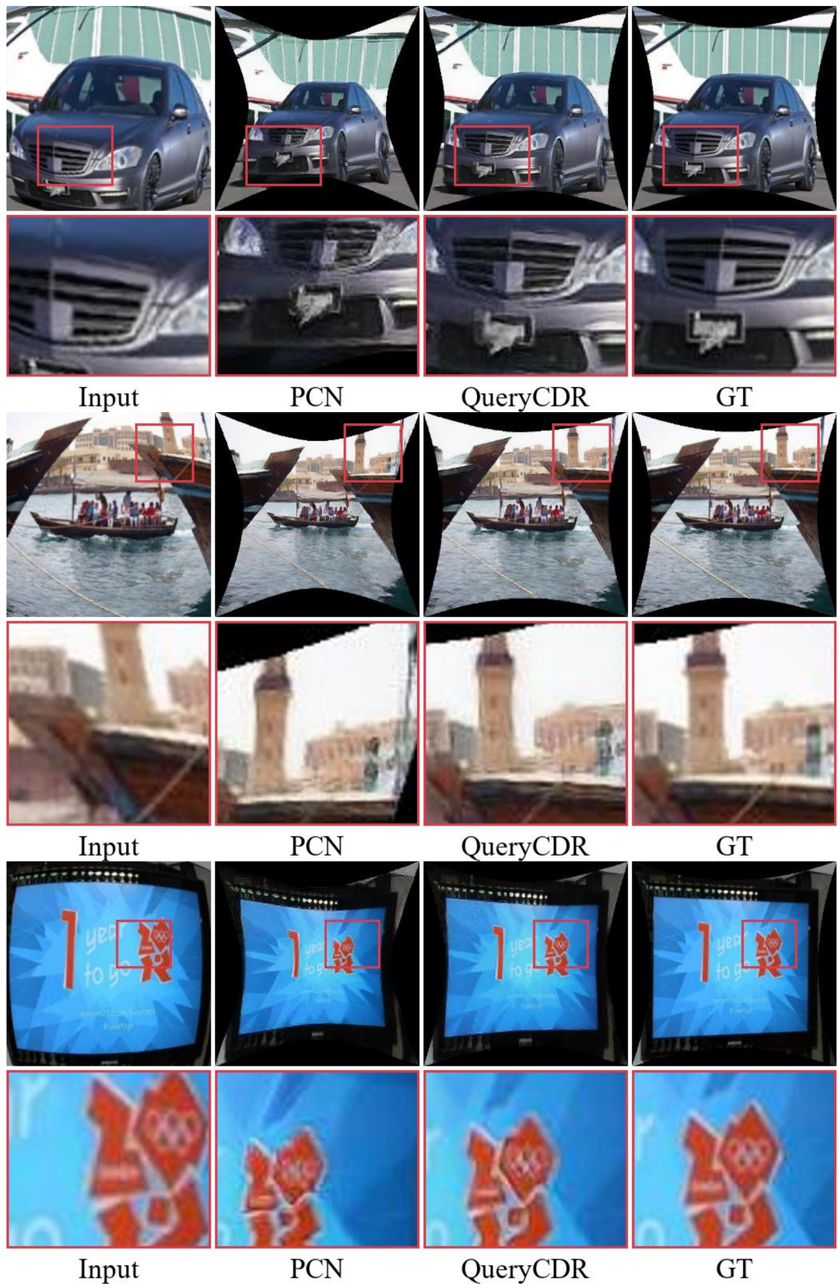

Fig. 5: Comparisons with the second best approach.

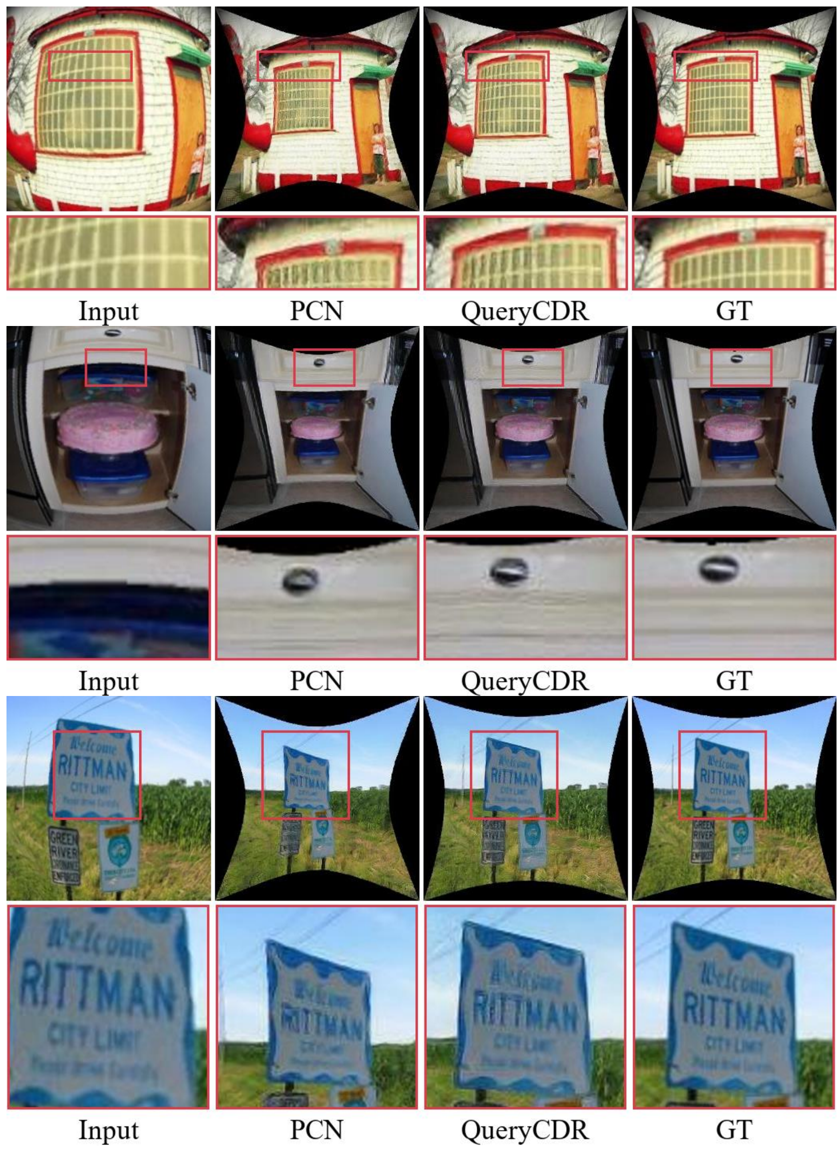

Fig. 6: Comparisons with the second best approach.