Supplementary Material for Sampling Innovation-Based Adaptive Compressive Sensing

Zhifu Tian, Tao Hu^{*}, Chaoyang Niu, Di Wu, Shu Wang Information Engineering University, Zhengzhou, China

tzhifu@qq.com, hutaoengineering@163.com, niucy2017@outlook.com, wudipaper@sina.com, shu1008@mail.ustc.edu.cn

In the Supplementary Material, we supplement the implementation details in Sec. 1. We further provide more experiments and analyses, including network convergence in Sec. 2, more comparisons with the state-of-the-art methods in Sec. 3, analysis of reconstruction errors in Sec. 4, effectiveness analysis of the innovation-guided multi-stage ASM in Sec. 5, complexity analysis of the lightweight reconstruction network in Sec. 6, ablation studies on the loss function in Sec. 7, and noise robustness in Sec. 8.

1. Implementation Details

For model training, following [6-8, 22, 26], we use a training set composed of 128×128 sub-images, which are randomly cropped from 40,000 unlabeled images from the COCO2017 dataset [12]. The sampling rate SR during training varies from 0.10 to 0.50. We employ the Adam [9] optimizer for model optimization and set the batch size to 16. The initial learning rate for the reconstruction network training of SIB-ACS is set to 1×10^{-4} . If there is no improvement in the aggregate PSNR at sampling rates of 0.10, 0.25, and 0.50 on the Set11 dataset [10] after five consecutive training epochs, we halve the learning rate. For the training of the lightweight reconstruction network in IE, each lightweight reconstruction network is co-trained with the sampling matrix. After completing the training of one stage of the lightweight reconstruction network and the sampling matrix, we freeze the parameters and then train the subsequent stages of the lightweight reconstruction network and sampling matrix.

For model evaluation, we employ widely used metrics such as PSNR and SSIM for our evaluation experiments, which are conducted on the widely used BSD68 [13] and Urban100 [11] datasets. For color images, evaluations are performed on the Y channel in the YCbCr color space. All experiments are executed on an NVIDIA Quadro RTX 6000 GPU, within a PyTorch-1.11.0 environment.



Figure 1. Convergence curves of PC-Net, CD-Net, and PCCD-Net on Set11 [10] at a sampling rate of 0.10.

2. Network Convergence

31.5

30.5

We investigate the convergence of PC-Net, CD-Net, and PCCD-Net over iterative phases on the Set11 dataset [10] with a sampling rate of 0.10, as shown in Fig. 1. The results demonstrate that the final convergence performance of PC-Net and CD-Net is inferior to that of PCCD-Net. This confirms the superiority of PCCD-Net over PC-Net and CD-Net, as discussed in Sec. 4.4 of the main text, rather than any reduction in model size caused by breakdown experiments. Moreover, PCCD-Net achieves convergence at the 24th phase. Therefore, we set the number of iterative phases K for PCCD-Net to 24.

3. More Comparisons with the state-of-the-art Methods

We compare the proposed SIB-ACS with eight other CS methods, including ISTA-Net⁺ [24], CSNet⁺ [16], COAST [23], TCS-Net [5], CSformer [22], SODAS-Net [17], AutoBCS [4], and SCNet [2], under five sampling rates ranging from low to high on the widely used BSD68 [13] and Urban100 [11] datasets, as shown in Tab. 1. CSNet⁺ and AutoBCS are purely DL methods based on pure CNN, TCS-Net and CSformer are pure DL methods based on Attention. ISTA-Net⁺, COAST, and SODAS-Net are DUNs, while SCNet is a self-supervised DUN. Despite the fact that the proposed SIB-ACS uses an adaptive CS

^{*}Corresponding authors.

Table 1. Average PSNR (dB) and SSIM results of other CS methods on BSD68 [13] and Urban100 [11] with different sampling ratios. The best results are marked in bold.



Figure 2. Visual comparisons of reconstruction errors on test003 from BSD68 [13] at the sampling ratio of 0.10 and imag048 from Urban100 [11] at a sampling ratio of 0.25. The best and second-best results are marked in red and blue colors, respectively.

model to handle adaptive image reconstruction at any sampling rate from 0.10 to 0.50, whereas other methods typically use a single model focused on image reconstruction at a specific sampling rate, the results in Tab. 1 demonstrate that the proposed SIB-ACS significantly outperforms existing methods in terms of image pixels and textures across all sampling rates.



Figure 3. Visual comparisons of reconstruction errors between US and AS on test033 from BSD68 [13] at a sampling ratio of 0.25.

Figure 4. Visual comparisons of reconstruction errors in different AS methods on img063 from Urban100 [11] at a sampling ratio of 0.25.

4. Analysis of Reconstruction Errors

Comparison of Reconstruction Errors with State-of-theart CS Methods. We supplement Sec. 4.2 of the main text with a comparison of reconstruction errors against 14 state-of-the-art CS methods, as depicted in Fig. 2. These 14 state-of-the-art CS methods include AMP-Net-9-BM [27]. MADUN [18], TranCS [15], FSOINet [3], DGUNet⁺ [14], LTwIST [6], DPC-DUN [19], OCTUF [20], NesTD-Net [7], UFC-Net [21], CPP-Net [8], CASNet [1], AMS-Net [25], and Uformer-ICS [26]. AMP-Net-9-BM, MADUN, TranCS, FSOINet, DGUNet⁺, LTwIST, DPC-DUN, OCTUF, NesTD-Net, UFC-Net, and CPP-Net are UCS methods, while CASNet, AMS-Net, and Uformer-ICS are ACS methods. Specifically, CASNet and AMS-Net have access to ground truth, while Uformer-ICS does not. The visual results in Fig. 2 show that the proposed SIB-ACS outperforms both UCS and other ACS methods in terms of overall image reconstruction error and error in challenging areas, under various sampling rates. The results suggest that the proposed SIB-ACS can efficiently allocate more sampling to challenging-to-reconstruct areas, thereby reducing the overall error of the reconstructed image.

Analysis of Reconstruction Errors in AS. We enrich Sec. 4.3 of the main text with a comparative analysis of reconstruction errors between uniform and adaptive sampling methods, as depicted in Fig. 3. The results demonstrate that the proposed AS method surpasses US methods in terms of effective sampling allocation. The proposed AS method significantly reduces the reconstruction error in areas that are challenging to reconstruct and decreases the overall image reconstruction error, thereby facilitating high-fidelity scene perception.

Analysis of Reconstruction Errors in Different AS Methods. We extend Sec. 4.3 of the main text with a comparative analysis of reconstruction errors among different AS methods, as depicted in Fig. 4. The reconstruction errors clearly show that the measurement errors-guided methods result in unstable AS, and the saliency-guided methods fail to correct AS due to the accumulation of initial sampling allocation over multiple stages. Both scenarios lead to the accumulation of anomalous sampling, resulting in uneven residual reconstruction errors. However, the proposed innovation-guided AS method, which minimizes the overall reconstruction error and rectifies AS through multi-stage feedback, ultimately results in a lower overall reconstruction error.

5. Effectiveness Analysis of the Innovation-Guided Multi-Stage ASM

Effectiveness Analysis of Innovation-Guided AS Methods. To analyze the effectiveness of the innovation-guided AS when applied to blocks of varying complexity within the same image, we examine the trend of reconstruction Mean Squared Error (MSE) with sampling rates in four blocks of different complexities in the cameraman image from the Set11 dataset [10], as depicted in Fig. 5. Based on the characteristics of innovation, the change in MSE with sampling rates, which is represented by the slope of the curve in Fig. 5, reflects the magnitude of the innovation. Fig. 5 suggests that the degree of reduction in reconstruction error is directly determined by the slope of the relationship curve between reconstruction error and sampling rates, rather than

Figure 5. The variation of reconstruction MSE for image blocks of different complexities in the cameramen image from Set11 [10] with the change in sampling rate.

Figure 6. Visual comparisons of ASA results, reconstructed images, and reconstruction errors of different AS methods between two-stage and multi-stage framework on test045 from BSD68 [13] at a sampling ratio of 0.10.

the recovery of reconstruction error resulting from historical sampling. Consequently, the innovation criterion can allocate sampling to areas with a more substantial reduction in reconstruction error more accurately than other criteria based on the image domain. Sec. 4.3 of the main text has already compared the performance of sampling innovation, measurement error, and saliency methods within a multi-stage framework. We further supplement the performance of sampling innovation, measurement error, and saliency methods within a two-stage framework, as depicted in Tab. 2 and Fig. 6(a). The results indicate that within the two-stage framework, the proposed sampling innovation method also surpasses other methods by a significant margin. The ASA of Fig. 6(a) demonstrates that the proposed sampling innovation method can avoid the influence of the recovered image components on the ASA judgment, thereby enabling it to more sensitively and extensively capture the recovery of image components brought about by the increase in sampling, ultimately achieving higher fidelity image reconstruction.

Effectiveness Analysis of a Multi-stage Framework. As illustrated in Fig.5, the degree of recovery from reconstruction error for different image blocks varies and continues

Table 2. PSNR (dB) and SSIM comparisons of different Adaptive Sampling (AS) methods in two-stage framework. The best results is marked in bold.

AS methods	BSI	D68	Urban100		
	SR = 0.10	SR = 0.25	SR = 0.10	SR = 0.25	
Measurement Error	28.40/0.8123	32.51/0.9130	28.65/0.8681	33.86/0.9431	
Saliency	28.45/0.8176	32.72/0.9177	28.48/0.8732	33.72/0.9460	
Sampling Innovation	29.29/0.8336	34.20/0.9296	28.95/0.8737	34.28/0.9467	

Table 3. Average PSNR (dB) and SSIM results of two-stage and multi-stage AS framework on Set11 [10] and Urban100 [11] with different CS ratios.

Framwork of ASM	BSI	D68	Urban100		
	SR = 0.25	SR = 0.40	SR = 0.25	SR = 0.40	
Two-stage	34.20/0.9296	38.19/0.9647	34.28/0.9467	37.80/0.9696	
Multi-stage	34.35/0.9312	38.38/0.9653	35.15/0.9516	38.93/0.9727	

to evolve with the increasing sampling rate. For instance, among the four image blocks, block 3 exhibits the largest recovery amount of reconstruction error with the increase of sampling rate within the range below 0.10, block 2 dominates when the rate ranges from 0.10 to 0.70, and block 4 takes the lead when the rate is above 0.70. Therefore, the ASA judgment under a single sampling rate situation is suboptimal, and a multi-stage feedback AS can yield a more accurate ASA. We conduct comparative experiments of innovation-guided two-stage AS and multi-stage AS, as presented in Tab. 3 and Fig. 6. For a fair comparison, the uniform sampling rate and adaptive sampling resources in the two-stage AS and multi-stage AS are kept identical. Tab. 3 reveals that the innovation-guided multi-stage AS framework significantly improves the quality of image reconstruction compared to the two-stage AS framework. Fig. 6 indicates that the innovation-guided multistage AS framework yields more reasonable ASA results than the two-stage AS framework. As illustrated in Fig. 6, the innovation-guided two-stage AS framework cannot correct the ASA error due to one-time ASA. However, the innovation-guided multi-stage AS framework, owing to the feedback from multi-stage AS, adjusts the ASA stage by stage, ultimately achieving a more precise ASA.

Table 4. Comprehensive comparisons of the model average PSNR, parameter size, and running time. The best results is marked in bold.

Models	OCTUF	NesTD-Net	UFC-Net	CPP-Net	CASNet	PC-Net8	PCCD-Net
PSNR (dB)	30.31	30.32	30.01	30.63	30.09	30.07	30.58
Params (M)	0.29	5.93	1.90	12.47	15.85	0.30	2.32
Time (s)	0.061	0.196	0.166	0.193	0.106	0.010	0.060
6000 5000 5000 5000 2000 2000 1000		+ Block 1 + Block 2 + Block 3 + Block 4 + Block 4 + Block 4 - Block 2 - Block 1 - Block 2 - Bloc	- PC-Nets - PC-Nets - PC-Nets - PCC-Nets - PCCD-Net - PCCD-Net - PCCD-Net	2000			Block 1 - PC-Net _y Block 2 - PC-Net _y Block 3 - PC-Net _y Block 4 - PC-Net _y Block 1 - PCCD-Net Block 3 - PCCD-Net Block 3 - PCCD-Net
0 0.01 0.0	12 0.03 0.04	0.05 0.06 0.07 0	.08 0.09 0.1	0 0.1	0.2 0.3 0.4	0.5 0.6 0 SR	.7 0.8 0.9 1

Figure 7. The variation of reconstruction MSE between PC-Net₈ and PCCD-Net for image blocks of different complexities in the cameramen image from Set11 [10] with the change in sampling rate.

Table 5. PSNR (dB) and SSIM comparisons of different loss functions on Set11 [10], BSD68 [13], and Urban100 [11] at a sampling ratio of 0.10. The best results is marked in bold.

Loss Function	Set11	BSD68	Urban100
l_1	32.31 /0.9154	29.57 /0.8343	29.59/0.8804
l_2	32.27/0.9159	29.57 /0.8351	29.65/0.8814
$l_1 + SSIM$	32.30/ 0.9177	29.54/ 0.8401	29.70/0.8859

6. Complexity Analysis of the Lightweight Reconstruction Network

The lightweight reconstruction network in the ASM is an 8-stage PC-Net (PC-Net₈). We evaluate the comprehensive performance of PC-Net₈, including average PSNR, model size, and running time, as detailed in Sec. 4.5 of the main text, as presented in Tab. 4. The results indicate that the model size of PC-Net₈ is close to that of the lightweight OCTUF [20], and PC-Net₈ has the shortest running time, thus ensuring the efficient operation of the ASM. Additionally, we examine the trend of MSE changes in the reconstruction of the four image blocks in Fig. 5 by PC-Net₈ and PCCD-Net under varying sampling rates, as illustrated in Fig. 7. The results reveal that although the image reconstruction quality of PC-Net₈ is lower than that of PCCD-Net, PC-Net₈ and PCCD-Net exhibit similar trends in the change of image reconstruction MSE with sampling rates, which demonstrates the feasibility and accuracy of employing PC-Net8 for IE. Considering the efficiency and accuracy of IE in ASM, we choose PC-Net₈ as the lightweight reconstruction network for IE.

7. Ablation of Loss Function

To assess the impact of different loss functions on model training, we independently utilize l_1 , l_2 , and $l_1 + SSIM$

Figure 8. Comparisons of noise robustness on BSD68 [13] at the sampling ratios of 0.10 and 0.25.

loss functions to train the model, and the PSNR and SSIM results on the Set11 [10], BSD68 [13], and Urban100 [11] datasets at a sampling rate of 0.10 are presented in the Tab. 5. The results suggest that the $l_1 + SSIM$ loss function generally provides superior reconstruction quality, and consistently outperforms in terms of SSIM across all tested scenarios. Considering both the pixel-level detail and texture of the reconstructed images, we select $l_1 + SSIM$ as the loss function for model training.

8. Noise Robustness

In practical applications, scene sensing is often affected by noise. We evaluate the reconstruction performance of SIB-ACS and other state-of-the-art methods in the presence of Gaussian noise with a mean of 0 and standard deviations of 0, 2, 4, and 6 on the BSD68 [13] dataset. The results at sampling rates of 0.10 and 0.25 are presented in Fig. 8. The results demonstrate that the proposed SIB-ACS outperforms other advanced methods under different sampling rates and varying degrees of noise, thus showcasing the excellent noise robustness of the proposed SIB-ACS.

References

- Bin Chen and Jian Zhang. Content-aware scalable deep compressed sensing. *IEEE Transactions on Image Processing*, 31:5412–5426, 2022. 3
- [2] Bin Chen, Xuanyu Zhang, Shuai Liu, Yongbing Zhang, and Jian Zhang. Self-supervised scalable deep compressed sensing. *International Journal of Computer Vision*, pages 1–36, 2024. 1, 2
- [3] Wenjun Chen, Chunling Yang, and Xin Yang. Fsoinet: feature-space optimization-inspired network for image compressive sensing. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), 2022. 3
- [4] Hongping Gan, Yang Gao, Chunyi Liu, Haiwei Chen, Tao Zhang, and Feng Liu. Autobcs: Block-based image compressive sensing with data-driven acquisition and noniterative reconstruction. *IEEE Transactions on Cybernetics*, 53 (4):2558–2571, 2021. 1, 2
- [5] Hongping Gan, Minghe Shen, Yi Hua, Chunyan Ma, and Tao Zhang. From patch to pixel: A transformer-based hierarchi-

cal framework for compressive image sensing. *IEEE Trans*actions on Computational Imaging, 9:133–146, 2023. 1, 2

- [6] Hongping Gan, Xiaoyang Wang, Lijun He, and Jie Liu. Learned two-step iterative shrinkage thresholding algorithm for deep compressive sensing. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023. 1, 3
- [7] Hongping Gan, Zhen Guo, and Feng Liu. Nestd-net: Deep nesta-inspired unfolding network with dual-path deblocking structure for image compressive sensing. *IEEE Transactions* on *Image Processing*, 2024. 3
- [8] Zhen Guo and Hongping Gan. Cpp-net: Embracing multiscale feature fusion into deep unfolding cp-ppa network for compressive sensing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 25086–25095, 2024. 1, 3
- [9] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2015. 1
- [10] Kuldeep Kulkarni, Suhas Lohit, Pavan Turaga, Ronan Kerviche, and Amit Ashok. Reconnet: Non-iterative reconstruction of images from compressively sensed measurements. In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, pages 449–458, 2016. 1, 3, 4, 5
- [11] Wei-Sheng Lai, Jia-Bin Huang, Narendra Ahuja, and Ming-Hsuan Yang. Fast and accurate image super-resolution with deep laplacian pyramid networks. *IEEE transactions on pattern analysis and machine intelligence*, 41(11):2599–2613, 2018. 1, 2, 3, 4, 5
- [12] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13, pages 740–755. Springer, 2014. 1
- [13] David Martin, Charless Fowlkes, Doron Tal, and Jitendra Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, pages 416–423. IEEE, 2001. 1, 2, 3, 4, 5
- [14] Chong Mou, Qian Wang, and Jian Zhang. Deep generalized unfolding networks for image restoration. In *Proceedings of* the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022. 3
- [15] Minghe Shen, Hongping Gan, Chao Ning, Yi Hua, and Tao Zhang. Transcs: a transformer-based hybrid architecture for image compressed sensing. *IEEE Transactions on Image Processing*, 31:6991–7005, 2022. 3
- [16] Wuzhen Shi, Feng Jiang, Shaohui Liu, and Debin Zhao. Image compressed sensing using convolutional neural network. *IEEE Transactions on Image Processing*, 29:375–388, 2019.
 1, 2
- [17] Jiechong Song and Jian Zhang. Sodas-net: Sideinformation-aided deep adaptive shrinkage network for compressive sensing. *IEEE Transactions on Instrumentation and Measurement*, 72:1–12, 2023. 1, 2
- [18] Jiechong Song, Bin Chen, and Jian Zhang. Memoryaugmented deep unfolding network for compressive sensing.

In Proceedings of the 29th ACM international conference on multimedia (ACM MM), pages 4249–4258, 2021. 3

- [19] Jiechong Song, Bin Chen, and Jian Zhang. Dynamic pathcontrollable deep unfolding network for compressive sensing. *IEEE Transactions on Image Processing*, 32:2202– 2214, 2023. 3
- [20] Jiechong Song, Chong Mou, Shiqi Wang, Siwei Ma, and Jian Zhang. Optimization-inspired cross-attention transformer for compressive sensing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (CVPR), pages 6174–6184, 2023. 3, 5
- [21] Xiaoyang Wang and Hongping Gan. Ufc-net: Unrolling fixed-point continuous network for deep compressive sensing. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 25149–25159, 2024. 3
- [22] Dongjie Ye, Zhangkai Ni, Hanli Wang, Jian Zhang, Shiqi Wang, and Sam Kwong. Csformer: Bridging convolution and transformer for compressive sensing. *IEEE Transactions* on Image Processing, 32:2827–2842, 2023. 1, 2
- [23] Di You, Jian Zhang, Jingfen Xie, Bin Chen, and Siwei Ma. Coast: Controllable arbitrary-sampling network for compressive sensing. *IEEE Transactions on Image Processing*, 30:6066–6080, 2021. 1, 2
- [24] Jian Zhang and Bernard Ghanem. Ista-net: Interpretable optimization-inspired deep network for image compressive sensing. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pages 1828–1837, 2018. 1, 2
- [25] Kuiyuan Zhang, Zhongyun Hua, Yuanman Li, Yongyong Chen, and Yicong Zhou. Ams-net: Adaptive multi-scale network for image compressive sensing. *IEEE Transactions on Multimedia*, 25:5676–5689, 2023. 3
- [26] Kuiyuan Zhang, Zhongyun Hua, Yuanman Li, Yushu Zhang, and Yicong Zhou. Uformer-ics: A u-shaped transformer for image compressive sensing service. *IEEE Transactions on Services Computing*, 2023. 1, 3
- [27] Zhonghao Zhang, Yipeng Liu, Jiani Liu, Fei Wen, and Ce Zhu. Amp-net: Denoising-based deep unfolding for compressive image sensing. *IEEE Transactions on Image Processing*, 30:1487–1500, 2020. 3