Image Restoration via Data-dependent Proximal Averaged Optimization

Pan Mu, Jian Chen, Risheng Liu*, Wei Zhong, Xin Fan, Zhongxuan Luo

Dalian University of Technology, Dalian, China

Team web: http://dutmedia.org/

Contents

1 The Proposed DPA Method

Problem

$$\mathbf{y} = \mathrm{K}(\mathbf{x}, \mathbf{n})$$

Observed

Latent

Noise

1 The Proposed DPA Method

Problem

Image deblurring

$$\mathbf{y} = \mathrm{K}(\mathbf{x}, \mathbf{n}) \quad \left\{ \; \mathbf{y} = \mathbf{k} \otimes \mathbf{x} + \mathbf{n} \right.$$

Observed ($\mathbf{y}$), Latent ($\mathbf{x}$), Noise ($\mathbf{n}$)

1 The Proposed DPA Method

Problem

Image deblurring

$$\mathbf{y} = \mathrm{K}(\mathbf{x}, \mathbf{n}) \begin{cases} \mathbf{y} = \mathbf{k} \otimes \mathbf{x} + \mathbf{n} \\ \mathbf{y} = \mathbf{x} + \mathbf{x}_r \end{cases}$$

Observed

Latent

Noise

Rain streaks removal

1 The Proposed DPA Method

Model

General MAP Model

$$p(\mathbf{x} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \mathbf{x})\, p(\mathbf{x})$$

$$\min_{\mathbf{x}} F(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$$

$f(\mathbf{x})$: Data fidelity

$g(\mathbf{x})$: Prior/Regularization

1 The Proposed DPA Method

Model

General MAP Model

$$p(\mathbf{x} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \mathbf{x})\, p(\mathbf{x})$$

$$\min_{\mathbf{x}} F(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$$

$f(\mathbf{x})$: Data fidelity

$g(\mathbf{x})$: Prior/Regularization

$$\mathbf{x} \in \mathbb{R}^{D}, \text{or}, \mathrm{h}(\mathbf{x}) \leq C$$

1 The Proposed DPA Method

Model

General MAP Model

$$p(\mathbf{x}\,|\,\mathbf{y}) \propto p(\mathbf{y}\,|\,\mathbf{x})\,p(\mathbf{x})$$

$$\min_{\mathbf{x}} F(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$$

Data fidelity

Prior/Regularization

$$\mathbf{x} \in \mathbb{R}^D, \mathrm{or}, \mathrm{h}(\mathbf{x}) \leq C$$

Energy based constraint

1 The Proposed DPA Method

Model

General MAP Model

$$p(\mathbf{x}\,|\,\mathbf{y}) \propto p(\mathbf{y}\,|\,\mathbf{x})\,p(\mathbf{x})$$

$$\min_{\mathbf{x}} F(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$$

- $f(\mathbf{x})$: Data fidelity
- $g(\mathbf{x})$: Prior/Regularization

$$\mathbf{x} \in \mathbb{R}^D, \text{or}, \mathrm{h}(\mathbf{x}) \le C$$

Energy based constraint

$$\mathbf{x} \in \arg\min_{\mathbf{x}} H(\mathbf{x}),$$

Ours

$$\min_{\mathbf{x}} F(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$$

$$\text{s.t.}, \mathbf{x} \in \arg\min_{\mathbf{x}} H(\mathbf{x}),$$

1 The Proposed DPA Method

Model

General MAP Model

$$p(\mathbf{x}\,|\,\mathbf{y}) \propto p(\mathbf{y}\,|\,\mathbf{x})\,p(\mathbf{x})$$

$$\min_{\mathbf{x}} F(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$$

Data fidelity

Prior/Regularization

$$\mathbf{x} \in \mathbb{R}^D, \text{or}, \mathrm{h}(\mathbf{x}) \leq C$$

Energy based constraint

$$\mathbf{x} \in \arg\min_{\mathbf{x}} H(\mathbf{x}),$$

Ours

$$\min_{\mathbf{x}} F(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$$

$$\text{s.t.}, \mathbf{x} \in \arg\min_{\mathbf{x}} H(\mathbf{x}),$$

1 The Proposed DPA Method

Algorithm

$$\min_{\mathbf{x}} F(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$$

$$\text{s.t.}, \mathbf{x} \in \arg\min_{\mathbf{x}} H(\mathbf{x}),$$

$$\mathcal{P}(\mathbf{x}^k, t_k) \in \arg\min_{\mathbf{x}} g(\mathbf{x}) + 1/(2t_k)\|\mathbf{x} - (\mathbf{x}^k - t_k \nabla f(\mathbf{x}^k))\|^2,$$

$$\mathcal{L}(\mathbf{x}^k, \tilde{\mathbf{x}}^k, \mu_k) \in \arg\min_{\mathbf{x}} H(\mathbf{x}) + \mu_k \|\mathbf{x} - \tilde{\mathbf{x}}^k\|^2.$$

1 The Developed DPA Method

Algorithm

$$\min_{\mathbf{x}} F(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$$

$$\mathcal{P}(\mathbf{x}^k, t_k) \in \arg\min_{\mathbf{x}} g(\mathbf{x}) + 1/(2t_k)\|\mathbf{x} - (\mathbf{x}^k - t_k \nabla f(\mathbf{x}^k))\|^2,$$

$$\text{s.t.}, \mathbf{x} \in \arg\min_{\mathbf{x}} H(\mathbf{x}),$$

$$\mathcal{L}(\mathbf{x}^k, \tilde{\mathbf{x}}^k, \mu_k) \in \arg\min_{\mathbf{x}} H(\mathbf{x}) + \mu_k \|\mathbf{x} - \tilde{\mathbf{x}}^k\|^2.$$

Averaged scheme

$$\mathcal{A}(\mathcal{P}, \mathcal{H}, \alpha^k) = \alpha^k \mathcal{P}(\mathbf{x}^k, t_k) + (1 - \alpha^k)\mathcal{L}(\mathbf{x}^k, \tilde{\mathbf{x}}^k, \mu_k).$$

Monotonous Correction

$$\mathcal{M}(\mathbf{u}^k, \mathbf{z}^k) = \begin{cases} \mathbf{u}^k, & \text{if } F(\mathbf{u}^k) < F(\mathbf{z}^k), \\ \mathbf{z}^k, & \text{else.} \end{cases}$$

1 The Proposed DPA Method

Algorithm

Algorithm 1 Data-dependent Proximal Average Framework

Require: The input $\mathbf{x}^0$, μ^0, α^0, parameters $t_k \in (0, 1/L^f)$.

```
while not converged do
    u^k = P(x^k, t_k),
    v^k = L(x^k, D^k, μ^k),
    z^k = A(u^k, v^k, α^k).
    x^{k+1} = M(u^k, z^k).
end while
```

$$\mathcal{P}(\mathbf{x}^k, t_k) \in \arg\min_{\mathbf{x}} g(\mathbf{x}) + 1/(2t_k)\|\mathbf{x} - (\mathbf{x}^k - t_k \nabla f(\mathbf{x}^k))\|^2,$$

$$\mathcal{L}(\mathbf{x}^k, \tilde{\mathbf{x}}^k, \mu_k) \in \arg\min_{\mathbf{x}} H(\mathbf{x}) + \mu_k \|\mathbf{x} - \tilde{\mathbf{x}}^k\|^2.$$

$$\mathcal{A}(\mathcal{P}, \mathcal{H}, \alpha^k) = \alpha^k \mathcal{P}(\mathbf{x}^k, t_k) + (1 - \alpha^k)\mathcal{L}(\mathbf{x}^k, \tilde{\mathbf{x}}^k, \mu_k).$$

$$\mathcal{M}(\mathbf{u}^k, \mathbf{z}^k) = \begin{cases} \mathbf{u}^k, & \text{if } F(\mathbf{u}^k) < F(\mathbf{z}^k), \\ \mathbf{z}^k, & \text{else.} \end{cases}$$

2 Applications and Discussion

Two Applications

	Deconvolution	Rain Streaks Removal
Task	$\mathbf{y} = \mathbf{k} \otimes \mathbf{x} + \mathbf{n}$	$\mathbf{y} = \mathbf{x} + \mathbf{x}_r$
f	$\|\mathbf{K}\mathbf{x} - \mathbf{y}\|^2$	$\|\mathbf{x} - \mathbf{D}^\top \boldsymbol{\alpha}\|^2 + \|\mathbf{x}_r - \mathbf{D}^\top \boldsymbol{\beta}\|^2$
g	$\rho\|\mathbf{W}\mathbf{x}\|_1$	$\rho_1\|\boldsymbol{\alpha}\|_1 + \rho_2\|\boldsymbol{\beta}\|_1 + \mathcal{I}_\Omega(\mathbf{x}, \mathbf{x}_r)$
H	$\|\mathbf{k} \otimes \mathbf{x} - \mathbf{y}\|^2 + \lambda\|\nabla \mathbf{x}\|_p$	$\|\mathbf{y} - \mathbf{x} - \mathbf{x}_r\|^2 + \lambda_1\|\nabla \mathbf{x}\|_p$

Table 1. Summary of two IR tasks (deconvolution and rain streaks removal) and the functional structures.

2 Applications and Discussion

Discussion

(1) Illustrate the effectiveness of DPA in image deconvolution task.

(a) PSNR/SSIM	(b) 23.87/0.70	(c) 26.90/0.84	(d) 27.22/0.85
Blurry	$\mathcal{P}$	$\mathcal{L}$	DPA (Ours)

2 Applications and Discussion

Discussion

(2) Theory results

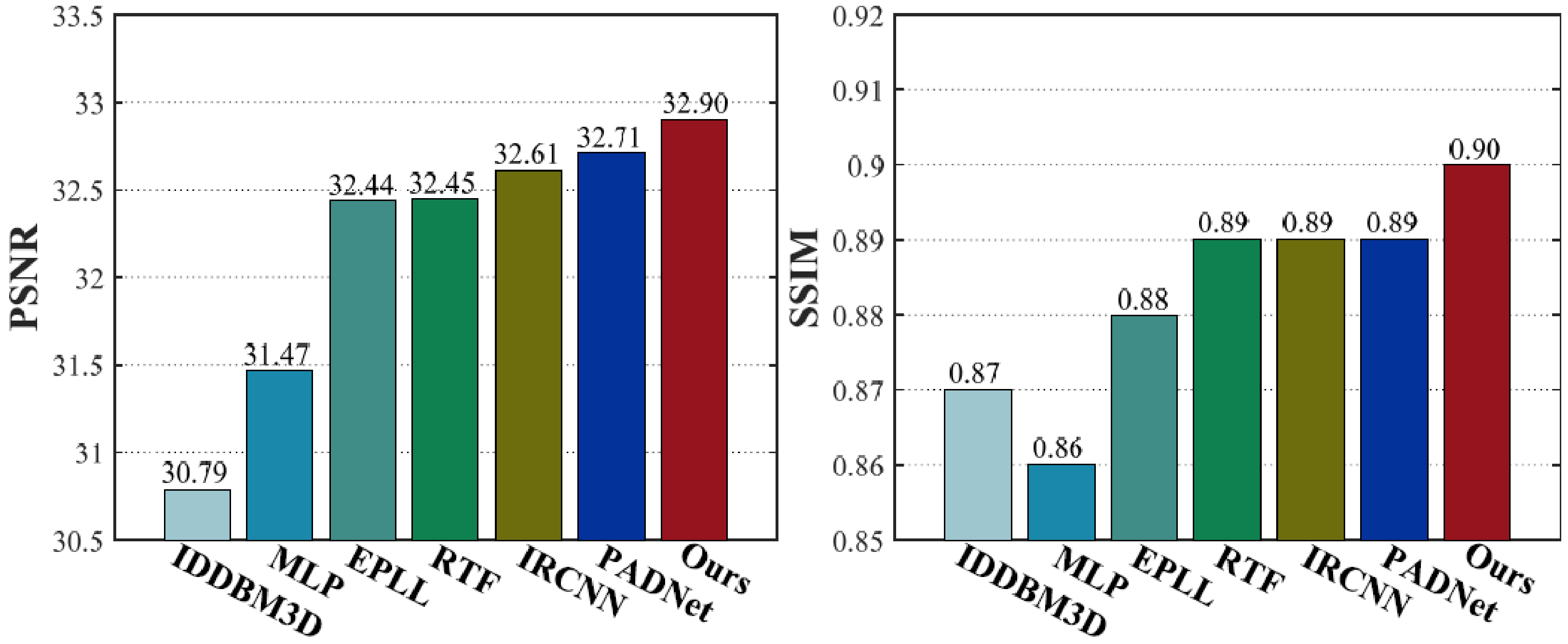
Sufficient Descent Property:

$$F(\mathbf{x}^{k+1}) \leq F(\mathbf{x}^k) - \left(1/(2t_k) - L^f/2\right)\|\mathbf{u}^k - \mathbf{x}^k\|^2.$$

There exists $\mathbf{x}^*$ be any accumulation of the sequence $\{\mathbf{x}^k\}$ satisfying $0 \in \partial F(\mathbf{x}^*)$. This means the proposed DPA convergence to a critical point of $\min_{\mathbf{x}} F(\mathbf{x})$.

3 Experiments

Image deconvolution



Averaged quantitative comparison of image deblurring on Sun et al. benchmark.

3 Experiments

Image deconvolution

–	28.73/0.8515	28.90/0.8665	30.27/0.8917	30.31/0.8943	30.53/0.9011
Blurry	IDDBM3D	RTF	IRCNN	PADNet	Ours

3 Experiments

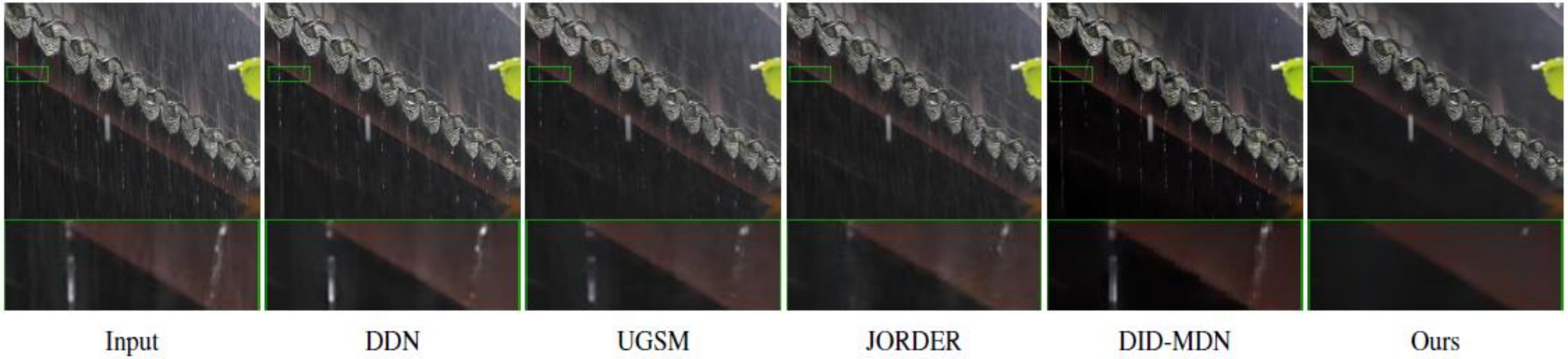
Rain streaks removal

- **Rain12:** including 12 synthesized rain images with only one type of rain streaks rendering technique;
- **Rain1400:** generated by 14 rainy images with different streak orientations and magnitudes
- **Rain100H:** collected from BSD200 and synthesized with five streak directions;

Methods	Rain12	Rain1400	Rain100H
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
LP	32.33 / 0.90	– / –	14.26 / 0.42
JCAS	31.61 / 0.92	26.80 / 0.85	15.23 / 0.52
DDN	33.41 / 0.94	29.99 / 0.89	17.93 / 0.57
UGSM	33.30 / 0.93	26.38 / 0.83	14.90 / 0.47
JORDER	35.93 / 0.95	28.90 / 0.90	23.45 / 0.75
DID-MDN	29.08 / 0.90	29.84 / 0.90	17.28 / 0.60
Ours	**36.39 / 0.96**	**31.33 / 0.91**	**24.30 / 0.80**

3 Experiments

Rain streaks removal



4 Conclusions

- **First, we provide a new insight for solving IR problems.**
- **Second, the developed model takes advantages of different domain knowledges.**
- **Third, the propagation of DPA converges to a critical point of the objective.**

Thanks