Supplementary Materials for
"Effective Snapshot Compressive-spectral Imaging via Deep Denoising and Total Variation Priors"

Haiquan Qiu[1], Yao Wang[1,2]*, Deyu Meng[1,3]
[1]Xi'an Jiaotong University, Xi'an, China
[2]Shanghai Em-Data Technology Co., Ltd., China
[3]The Macau University of Science and Technology, Macau, China
qiuhaiquan@stu.xjtu.edu.cn, yao.s.wang@gmail.com, dymeng@mail.xjtu.edu.cn

1. Structure of this Document

This document is organized as follows. Section 2 provides some notations in this document. Some preliminaries are presented in Section 3. The proof of fixed point convergence is presented in Section 4. More experimental results are reported in Section 5.

2. Notations

For any $x, y \in \mathbb{R}^d$, $\langle x, y \rangle = x^T y$ denotes the inner product. In SCI problem, we can assume $d = nB$. Let $\mathbf{H} \in \mathbb{R}^{m \times d}$ and $\mathbf{y} \in \mathbb{R}^m$. We also assume $m = n$. $\mathbf{I}$ is identity mapping.

3. Preliminaries

An mapping $T : \mathbb{R}^d \to \mathbb{R}^d$ is L-Lipschitz if

$$\|T(x) - T(y)\| \leq L\|x - y\|$$

for any $x, y \in \mathbb{R}^d$. If T is L-Lipschitz with $L \leq 1$, we say T is nonexpansive. If T is L-Lipschitz with $L < 1$, we say T is a contraction. Next we shall give the assumptions used in the fixed point analysis.

Assumption 1 (Assumption (A) in [3]). *We assume that all denoisers $\mathcal{D}_\sigma : \mathbb{R}^d \mapsto \mathbb{R}^d$ used in our method satisfy*

$$\|(\mathcal{D}_\sigma - \mathbf{I})(x) - (\mathcal{D}_\sigma - \mathbf{I})(y)\|_2 \leq \epsilon \|x - y\|_2 \quad (1)$$

for all $x, y \in \mathbb{R}^d$ for some $\epsilon > 0$.

Assumption 2 (Assumption 1 in [4]). *Assume that $\{R_j\}_{j=1}^n > 0$ which means for each spatial location j, the B-frame modulation masks at this location have at least one non-zero entries. We further assume $R_{\max} > R_{\min}$.*

A mapping P projecting x onto linear manifold $\mathbf{y} = \mathbf{H}x$ can be expressed as

$$P(x) = x + \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}(\mathbf{y} - \mathbf{H}x).$$

Following lemmas are required to prove Theorem 1 and 2.

Lemma 1 (Lemma 2[1] in [4]). *For any $x \in \mathbb{R}^d, d = nB$, consider $\mathbf{H}$ satisfies Assumption 2, then*

$$\frac{R_{\min}}{R_{\max}}\|x\|^2 \leq \|\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}x\|^2 \leq \|x\|^2.$$

Lemma 2. *$\mathcal{S} = \{\mathcal{D}_\sigma : \sigma \in S\}$ is a set of denoiser satisfying Assumption 1 and $|\mathcal{S}| < \infty$. Then the weighted denoiser of $\mathcal{S}$:*

$$D_w(x) = \sum_{\sigma \in S} w_\sigma \mathcal{D}_\sigma(x)$$

also satisfies Assumption 1, where $\sum_{\sigma \in S} w_\sigma = 1, w_\sigma \geq 0, \forall \sigma \in S$.

Proof.

$$\begin{aligned}
&\|(D_w - \mathbf{I})(x) - (D_w - \mathbf{I})(y)\|_2 \\
=&\| \sum_{\sigma \in S} w_\sigma(\mathcal{D}_\sigma - \mathbf{I})(x) - \sum_{\sigma \in S} w_\sigma(\mathcal{D}_\sigma - \mathbf{I})(y)\|_2 \\
=&\| \sum_{\sigma \in S} w_\sigma[(\mathcal{D}_\sigma - \mathbf{I})(x) - (\mathcal{D}_\sigma - \mathbf{I})(y)]\|_2 \\
\leq& \sum_{\sigma \in S} w_\sigma\|(\mathcal{D}_\sigma - \mathbf{I})(x) - (\mathcal{D}_\sigma - \mathbf{I})(y)\|_2 \\
\leq& \sum_{\sigma \in S} w_\sigma \epsilon \|x - y\|_2 = \epsilon \|x - y\|_2,
\end{aligned}$$

where the first inequality follows from the triangle inequality. □

3.1. Equivalence of PnP-DRS and PnP-ADMM

In this section, we show the equivalence of PnP-DRS and PnP-ADMM, which is motivated from Section 9.1 in [3]. With the form of PnP-DRS, we substitute $z^{(k)} = x^{(k)} - u^{(k)}$ to have

$$
\begin{aligned}
x^{(k+1/2)} =& \mathcal{D}_\sigma(x^{(k)} - u^{(k)}), \\
x^{(k+1)} =& P\left(x^{(k+1/2)} + \left(u^{(k)} + (x^{(k+1/2)} - x^{(k)})\right)\right), \\
u^{(k+1)} =& u^{(k)} + (x^{(k+1/2)} - x^{(k)}).
\end{aligned}
$$

Reorder the iteration to get that

$$
\begin{aligned}
x^{(k+1/2)} =& \mathcal{D}_\sigma(x^{(k)} - u^{(k)}), \\
u^{(k+1)} =& u^{(k)} + (x^{(k+1/2)} - x^{(k)}), \\
x^{(k+1)} =& P\left(x^{(k+1/2)} + u^{(k+1)}\right).
\end{aligned}
$$

We rewrite $\tilde{y}^{(k+1)} = x^{k+1/2}$ and $\tilde{x}^{(k+1)} = x^{(k)}$, thus

$$
\begin{aligned}
\tilde{x}^{(k+1)} =& P\left(\tilde{y}^{(k)} + u^{(k)}\right), \\
\tilde{y}^{(k+1)} =& \mathcal{D}_\sigma(\tilde{x}^{(k+1)} - u^{(k)}), \\
u^{(k+1)} =& u^{(k)} + (\tilde{y}^{(k+1)} - \tilde{x}^{(k+1)}).
\end{aligned}
$$

Finally, we get PnP-ADMM.

3.2. Equivalence of PnP-ADMM and Accelerated PnP-GAP

In this section, we show that the PnP-ADMM is equivalent to accelerated PnP-GAP. We are going to convert the following PnP-ADMM into accelerated PnP-GAP. Precisely,

$$
\mathbf{x}^{(k+1)} = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{x} - \mathbf{v}^{(k)} - \mathbf{u}^{(k)}\|_2^2, \text{ subject to } \mathbf{H}\mathbf{x} = \mathbf{y}, \tag{2a}
$$

$$
= P(\mathbf{v}^{(k)} + \mathbf{u}^{(k)}), \tag{2b}
$$

$$
\mathbf{v}^{(k+1)} = \arg\min_{\mathbf{v}} \lambda g(\mathbf{v}) + \frac{\gamma}{2}\|\mathbf{x}^{(k+1)} - \mathbf{v} - \mathbf{u}^{(k)}\|_2^2 \tag{2c}
$$

$$
= \mathcal{D}_\sigma(\mathbf{x}^{(k+1)} - \mathbf{u}^{(k)}), \tag{2d}
$$

$$
\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + (\mathbf{v}^{(k+1)} - \mathbf{x}^{(k+1)}). \tag{2e}
$$

We can rewrite $\tilde{\mathbf{x}}^{(k)} = \mathbf{x}^{(k)} - \mathbf{u}^{(k-1)}$, then (2d) is converted to

$$
\tilde{\mathbf{x}}^{(k+1)} = \mathbf{v}^{(k)} + \mathbf{H}(\mathbf{H}\mathbf{H}^T)^{-1}\left(\mathbf{y}^{(k)} - \mathbf{H}\mathbf{v}^{(k)}\right),
$$

where $\mathbf{y}^{(k)} = \mathbf{y} - \mathbf{H}\mathbf{u}^{(k)}$. Using (2e), $\mathbf{y}^{(k+1)}$ is updated with

$$
\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} + \mathbf{y} - H\mathbf{v}^{(k+1)}.
$$

Now the PnP-ADMM (2) is converted into

$$
\tilde{\mathbf{x}}^{(k+1)} = \mathbf{v}^{(k)} + \mathbf{H}(\mathbf{H}\mathbf{H}^T)^{-1}\left(\mathbf{y}^{(k)} - \mathbf{H}\mathbf{v}^{(k)}\right), \tag{3}
$$

$$
\mathbf{v}^{(k+1)} = \mathcal{D}_\sigma(\tilde{\mathbf{x}}^{(k+1)}), \tag{4}
$$

$$
\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} + \mathbf{y} - \mathbf{H}\mathbf{v}^{(k+1)}. \tag{5}
$$

We then reorder the sequence to get that

$$
\begin{aligned}
\mathbf{y}^{(k+1)} =& \mathbf{y}^{(k)} + \mathbf{y} - \mathbf{H}\mathbf{v}^{(k+1)}, \\
\tilde{\mathbf{x}}^{(k+2)} =& \mathbf{v}^{(k+1)} + \mathbf{H}(\mathbf{H}\mathbf{H}^T)^{-1}\left(\mathbf{y}^{(k+1)} - \mathbf{H}\mathbf{v}^{(k+1)}\right), \\
\mathbf{v}^{(k+2)} =& \mathcal{D}_\sigma(\tilde{\mathbf{x}}^{(k+2)}),
\end{aligned}
$$

which is the accelerated PnP-GAP used in our paper.

In Section 3.1 and Section 3.2, we convert the proof of convergence of accelerated PnP-GAP into the proof of convergence of PnP-DRS. The relationship between different algorithms is

$$
\begin{aligned}
&\text{Convergence of PnP-DRS} \\
\Rightarrow\ &\text{Convergence of PnP-ADMM} \\
\Rightarrow\ &\text{Convergence of accelerated PnP-GAP.}
\end{aligned}
$$

If we can prove PnP-DRS converges to a fixed point, it indicates accelerated PnP-GAP converges to a fixed point. Therefore, in the next section, our proof is focused on the convergence of PnP-DRS.

Before providing the proof, we have the following two important equations:

$$
\begin{aligned}
&\|\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}x\|^2 \\
=&x^T\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}x \\
=&x^T\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}x \\
=&\langle\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}x, x\rangle
\end{aligned} \tag{6}
$$

and

$$
\begin{aligned}
&\langle\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}x, \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}y\rangle \\
=&x^T\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}y \\
=&x^T\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}y \\
=&\langle\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}x, y\rangle \\
=&\langle\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}y, x\rangle
\end{aligned} \tag{7}
$$

for any $x, y \in \mathbb{R}^d$.

4. Fixed-point Convergence

Theorem 1. *Assume* $\mathbf{H}$ *satisfies Assumption 2. Then the following operator*

$$
G = \mathcal{D}_\sigma \circ P
$$

is a contraction if $\mathcal{D}_\sigma$ *satisfies Assumption 1 and*

$$
0 < \epsilon < \sqrt{\frac{R_{\max}}{R_{\max} - R_{\min}}} - 1.
$$

Proof. In order to prove G is a contraction, for any $x_1, x_2 \in \mathbb{R}^d$ we have

$$
\begin{aligned}
&\|G(x_1) - G(x_2)\| \\
=&\|\mathcal{D}_\sigma \circ P(x_1) - \mathcal{D}_\sigma \circ P(x_2)\| \\
=&\|(\mathcal{D}_\sigma - \mathbf{I}) \circ P(x_1) - (\mathcal{D}_\sigma - \mathbf{I}) \circ P(x_2) + P(x_1) - P(x_2)\| \\
\leq&(\epsilon + 1)\|P(x_1) - P(x_2)\|.
\end{aligned}
$$

Last inequality follows from Assumption 1 and triangle inequality. Because

$$
\begin{aligned}
&\|P(x_1) - P(x_2)\|^2 \\
=&\|(x_1 - x_2) - \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}(x_1 - x_2)\|^2 \\
=&\|x_1 - x_2\|^2 - \|\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}(x_1 - x_2)\|^2 \\
\leq&(1 - \frac{R_{\min}}{R_{\max}})\|x_1 - x_2\|^2,
\end{aligned}
$$

we can have

$$
\begin{aligned}
&\|G(x_1) - G(x_2)\| \\
\leq&(\epsilon + 1)\|P(x_1) - P(x_2)\| \\
\leq&(\epsilon + 1)\sqrt{1 - \frac{R_{\min}}{R_{\max}}}\|x_1 - x_2\| \\
<&\|x_1 - x_2\|.
\end{aligned}
$$

Then G is a contraction. $\square$

Theorem 2. *Assume $\mathbf{H}$ satisfies Assumption 2. Let P be a Euclidean projection on linear manifold $\mathbf{y} = \mathbf{H}x$. Then*

$$
T = \frac{1}{2}\mathbf{I} + \frac{1}{2}(2P - \mathbf{I})(2\mathcal{D}_\sigma - \mathbf{I})
$$

is a contraction if $\mathcal{D}_\sigma$ satisfies Assumption 1 and

$$
0 < \epsilon < 1 - \sqrt{1 - \frac{R_{\min}}{R_{\max}}}.
$$

Proof. From Section 3, we need to prove the convergence of PnP-DRS which has the form

$$
\begin{aligned}
T =&\frac{1}{2}\mathbf{I} + \frac{1}{2}(2P - \mathbf{I})(2\mathcal{D}_\sigma - \mathbf{I}) \\
=&\frac{1}{2}\mathbf{I} + \frac{1}{2}2P \circ (2\mathcal{D}_\sigma - \mathbf{I}) - \frac{1}{2}(2\mathcal{D}_\sigma - \mathbf{I}) \\
=&\mathbf{I} - \mathcal{D}_\sigma + P \circ (2\mathcal{D}_\sigma - \mathbf{I}).
\end{aligned}
$$

Because $P(x) = x + \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}(\mathbf{y} - \mathbf{H}x)$, let T operates on $x \in \mathbb{R}^d$:

$$
\begin{aligned}
&T(x) \\
=&x - \mathcal{D}_\sigma(x) + 2\mathcal{D}_\sigma(x) - x \\
&+ \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}[\mathbf{y} - \mathbf{H}(2\mathcal{D}_\sigma(x) - x)] \\
=&\mathcal{D}_\sigma(x) - \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}[2\mathcal{D}_\sigma(x) - x] + \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{y}.
\end{aligned}
$$

Then we denote $\mathbf{S} := \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}$. Using (6) and (7), we have the following equations:

$$
\|\mathbf{S}x\|^2 = \langle \mathbf{S}x, x \rangle
$$

and

$$
\langle \mathbf{S}x, \mathbf{S}y \rangle = \langle \mathbf{S}x, y \rangle = \langle \mathbf{S}y, x \rangle.
$$

We prove that T a contraction from the definition. For any $x_1, x_2 \in \mathbb{R}^d$, we have

$$
\begin{aligned}
&\|T(x_1) - T(x_2)\|^2 \\
=&\|\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2) \\
&- \mathbf{S}[2\mathcal{D}_\sigma(x_1) - 2\mathcal{D}_\sigma(x_2) - (x_1 - x_2)]\|^2 \\
=&\|\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)\|^2 \\
&+ \|\mathbf{S}[2\mathcal{D}_\sigma(x_1) - 2\mathcal{D}_\sigma(x_2) - (x_1 - x_2)]\|^2 \\
&- 2\langle \mathbf{S}[2\mathcal{D}_\sigma(x_1) - 2\mathcal{D}_\sigma(x_2) - (x_1 - x_2)], \\
&\quad \mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)\rangle. \quad (8)
\end{aligned}
$$

Next, we calculate terms in (8) to get

$$
\begin{aligned}
&\|\mathbf{S}[2\mathcal{D}_\sigma(x_1) - 2\mathcal{D}_\sigma(x_2) - (x_1 - x_2)]\|^2 \\
=&4\|\mathbf{S}[\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)]\|^2 + \|\mathbf{S}(x_1 - x_2)\|^2 \\
&- 4\langle \mathbf{S}[\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)], \mathbf{S}(x_1 - x_2)\rangle \\
=&4\|\mathbf{S}[\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)]\|^2 + \|\mathbf{S}[x_1 - x_2]\|^2 \\
&- 4\langle \mathbf{S}(x_1 - x_2), \mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)\rangle
\end{aligned}
$$

and

$$
\begin{aligned}
&-2\langle \mathbf{S}[2\mathcal{D}_\sigma(x_1) - 2\mathcal{D}_\sigma(x_2) - (x_1 - x_2)], \\
&\quad \mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)\rangle \\
=&-4\langle \mathbf{S}[\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)], \mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)\rangle \\
&+ 2\langle \mathbf{S}(x_1 - x_2), \mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)\rangle \\
=&-4\|\mathbf{S}[\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)]\|^2 \\
&+ 2\langle \mathbf{S}(x_1 - x_2), \mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)\rangle.
\end{aligned}
$$

Plugging these terms into (8), we further get

$$
\begin{aligned}
&\|T(x_1) - T(x_2)\|^2 \\
=&\|\mathbf{S}[2\mathcal{D}_\sigma(x_1) - 2\mathcal{D}_\sigma(x_2) - (x_1 - x_2)]\|^2 \\
&- 2\langle \mathbf{S}[2\mathcal{D}_\sigma(x_1) - 2\mathcal{D}_\sigma(x_2) - (x_1 - x_2)], \mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)\rangle \\
&+ \|\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)\|^2 \\
=&\|\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)\|^2 + 4\|\mathbf{S}[\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)]\|^2 \\
&+ \|\mathbf{S}(x_1 - x_2)\|^2 - 4\langle \mathbf{S}(x_1 - x_2), \mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)\rangle \\
&+ 2\langle \mathbf{S}(x_1 - x_2), \mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)\rangle \\
&- 4\|\mathbf{S}[\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)]\|^2 \\
=&\|\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)\|^2 + \|\mathbf{S}(x_1 - x_2)\|^2 \\
&- 2\langle \mathbf{S}(x_1 - x_2), \mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2)\rangle \\
=&\|\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2) - \mathbf{S}(x_1 - x_2)\|^2. \quad (9)
\end{aligned}
$$

Then we have

$$\|T(x_1) - T(x_2)\|$$
$$=\|\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2) - \mathbf{S}(x_1 - x_2)\|$$
$$=\|\mathcal{D}_\sigma(x_1) - \mathcal{D}_\sigma(x_2) - \mathbf{S}(x_1 - x_2) - (x_1 - x_2) + (x_1 - x_2)\|$$
$$=\|(\mathcal{D}_\sigma - \mathbf{I})(x_1) - (\mathcal{D}_\sigma - \mathbf{I})(x_2) + (\mathbf{I} - \mathbf{S})(x_1 - x_2)\|$$
$$\leq\|(\mathcal{D}_\sigma - \mathbf{I})(x_1) - (\mathcal{D}_\sigma - \mathbf{I})(x_2)\| + \|(\mathbf{I} - \mathbf{S})(x_1 - x_2)\|$$
$$\leq\epsilon\|x_1 - x_2\| + \|(\mathbf{I} - \mathbf{S})(x_1 - x_2)\|. \tag{10}$$

Thus we derive the bound of $\|(\mathbf{I} - S)(x_1 - x_2)\|$ as

$$\|(\mathbf{I} - \mathbf{S})(x_1 - x_2)\|^2$$
$$=\|x_1 - x_2 - \mathbf{S}(x_1 - x_2)\|^2$$
$$=\|x_1 - x_2\|^2 + \|\mathbf{S}(x_1 - x_2)\|^2 - 2\langle \mathbf{S}(x_1 - x_2), x_1 - x_2 \rangle$$
$$=\|x_1 - x_2\|^2 - \|\mathbf{S}(x_1 - x_2)\|^2$$
$$\leq(1 - \frac{R_{\min}}{R_{\max}})\|x_1 - x_2\|^2.$$

The last inequality follows from Lemma 1. Plugging this into (10) and using condition $\epsilon < 1 - \sqrt{1 - \frac{R_{\min}}{R_{\max}}}$, we get

$$\|T(x_1) - T(x_2)\|$$
$$\leq\epsilon\|x_1 - x_2\| + \|(\mathbf{I} - \mathbf{S})(x_1 - x_2)\|$$
$$\leq(\epsilon + \sqrt{1 - \frac{R_{\min}}{R_{\max}}})\|x_1 - x_2\|$$
$$<\|x_1 - x_2\|.$$

This means T is a contraction. Finally, PnP-DRS converges to a fixed point which is equivalent to the convergence of PnP-ADMM and accelerated PnP-GAP. $\square$

5. More Experiments

5.1. More Results of the Experiments in the Paper

This subsection shows more results of the experiment in our paper. Detailed results are shown in Figure 1, Figure 2, Figure 3 and Table 3.

5.2. Comparison with Learned Prior

In this subsection, we reported experiment results compared with learned prior method [1]. We use the four data (named 103, 101, 73, 92) provided in the code of [1][2]. The results are presented in Table 1.

5.3. Comparison with Deep Learning

In this section, we compared our method with one popular deep learning method, i.e., λ-Net [2]. We use a binary mask that differs from that used in the training phase for testing. λ-Net and our method have comparable results while it requires a couple of days for training. The results are reported in Table 2.

---
[2] https://github.com/KAIST-VCLAB/deepcassi

5.4. Results on Real Data: `object`

We conduct experiment on real data `object`. The results is listed in Table 4.

Table 4. Results on real scene: `object`

	2DTV	DeSCI	Ours (2DTV)
`object`	36.75, 0.8857	36.21, 0.8501	37.72, 0.8717

References

[1] Inchang Choi, Daniel S. Jeon, Giljoo Nam, Diego Gutierrez, and Min H. Kim. High-quality hyperspectral reconstruction using a spectral prior. *ACM Transactions on Graphics (Proc. SIGGRAPH Asia 2017)*, 36(6):218:1–13, 2017. 4, 6

[2] Xin Miao, Xin Yuan, Yunchen Pu, and Vassilis Athitsos. lambda-net: Reconstruct hyperspectral images from a snapshot measurement. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4058–4068. IEEE, 2019. 4

[3] Ernest K Ryu, Jialin Liu, Sicheng Wang, Xiaohan Chen, Zhangyang Wang, and Wotao Yin. Plug-and-play methods provably converge with properly trained denoisers. *arXiv preprint arXiv:1905.05406*, 2019. 1, 2

[4] Xin Yuan, Yang Liu, Jinli Suo, and Qionghai Dai. Plug-and-play algorithms for large-scale snapshot compressive imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1447–1457, 2020. 1
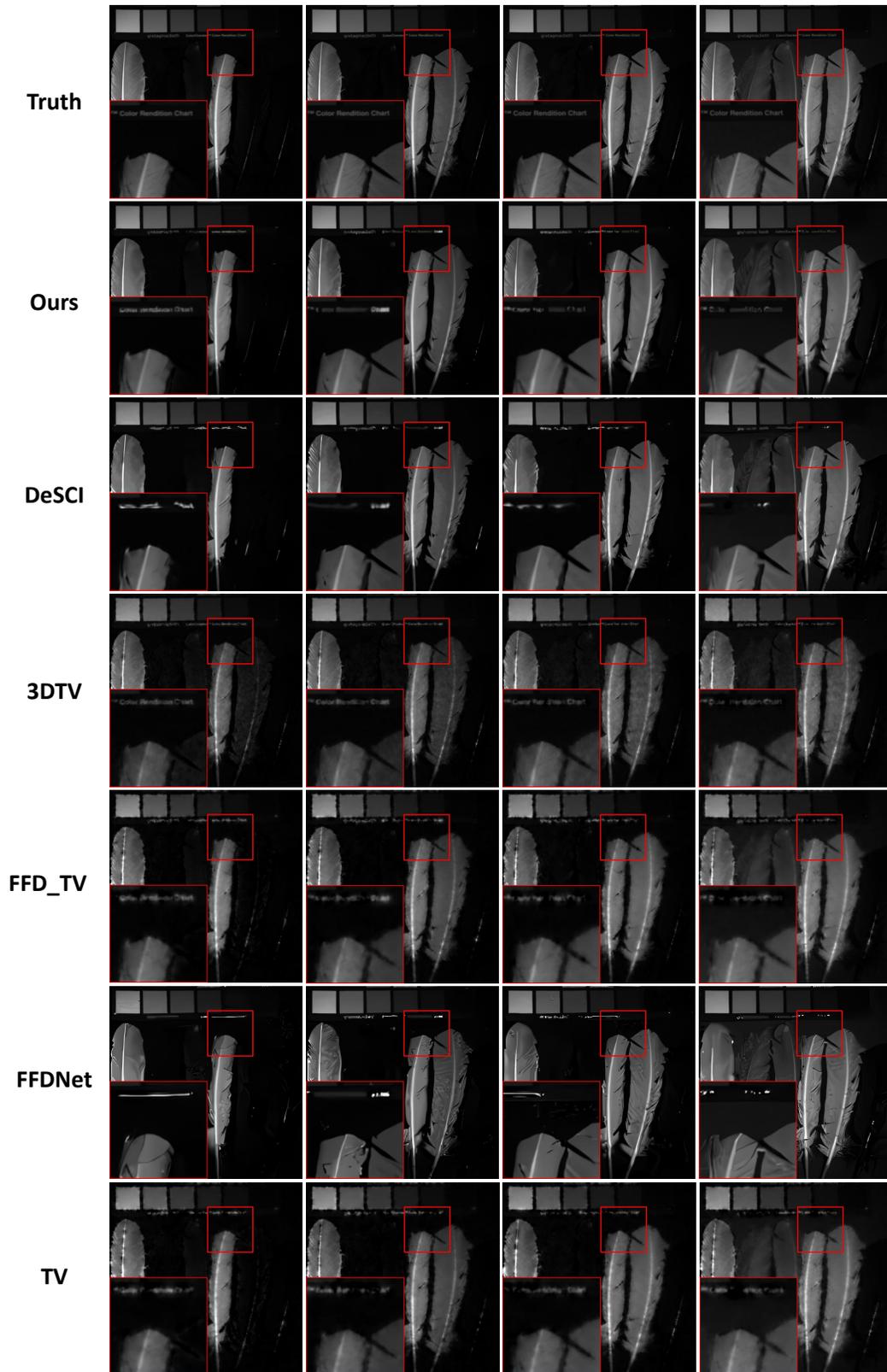
Figure 1. Simulated data: `feathers` from CAVE. The four frames are at wavelengths 580nm, 620nm, 660nm, and 700nm.

Table 1. Comparison with learned prior method AE [1].

	103	101	73	92	Average
AE	36.75, 0.9726	**38.35, 0.9694**	38.19, 0.9635	32.49, 0.8874	36.45, **0.9482**
3DTV	34.65, 0.9590	35.96, 0.9396	36.36, 0.9490	31.35, 0.8516	34.58, 0.9248
DeSCI	23.98, 0.8053	26.15, 0.8334	28.36, 0.8492	20.76, 0.6865	24.81, 0.7936
Our (3DTV)	**37.05, 0.9735**	38.14, 0.9599	**38.56, 0.9603**	**32.96, 0.8989**	**36.68, 0.9482**

Table 2. Comparison with deep learning method λ-Net.

	3DTV	2DTV	FFDNet	FFDNet-TV	DeSCI	λ-Net	Ours (3DTV)	Ours (2DTV)
Scene 1	31.83, 0.9179	34.68, 0.9321	27.91, 0.8794	35.08, 0.9237	36.07, 0.9505	37.99, 0.8971	37.12, 0.9626	38.16, 0.9630
Scene 2	22.38, 0.7542	27.39, 0.9165	22.55, 0.7463	27.39, 0.9114	30.64, 0.9580	32.70, 0.9465	30.11, 0.9558	30.11, 0.9566
Scene 3	28.66, 0.8851	28.52, 0.8888	21.75, 0.7236	28.04, 0.8883	29.87, 0.9132	34.02, 0.9524	31.70, 0.9156	31.38, 0.9040
Scene 4	25.44, 0.8205	31.72, 0.9309	24.97, 0.8055	32.47, 0.9303	40.35, 0.9780	30.11, 0.9247	31.89, 0.9404	38.25, 0.9691
Scene 5	31.05, 0.8782	31.63, 0.8599	26.17, 0.7643	32.02, 0.8547	33.86, 0.9038	38.10, 0.9330	34.73, 0.9313	34.70, 0.9265
Scene 6	26.47, 0.8517	28.04, 0.8613	26.11, 0.8725	28.79, 0.8525	33.59, 0.9421	30.73, 0.9222	31.44, 0.9234	32.82, 0.9326
Scene 7	29.20, 0.8841	33.90, 0.9287	24.81, 0.8128	34.31, 0.9279	35.76, 0.9515	37.15, 0.9675	34.70, 0.9431	36.17, 0.9476
Scene 8	26.94, 0.8716	30.12, 0.8779	21.64, 0.7117	30.29, 0.8703	31.34, 0.9061	34.35, 0.9454	30.12, 0.9086	31.62, 0.9044
Scene 9	33.31, 0.9350	35.31, 0.9569	37.64, 0.9485	35.90, 0.9541	40.87, 0.9694	36.04, 0.9264	37.07, 0.9705	40.56, 0.9746
Scene 10	25.23, 0.8307	27.59, 0.8431	20.37, 0.6289	27.83, 0.8400	28.96, 0.8746	29.47, 0.9062	28.81, 0.8803	28.99, 0.8731
Average	28.05, 0.8629	30.89, 0.8996	25.39, 0.7893	31.21, 0.8953	34.13, 0.9347	34.07, 0.9321	32.77, 0.9332	**34.28, 0.9352**

Table 3. The average result of PSNR in dB (left entry in each cell) and SSIM (right entry in each cell) by different algorithms on some data from CAVE.

Data	2DTV	3DTV	FFDNet	DeSCI	FFDNet-TV	Ours (2DTV)	Ours (3DTV)
feathers	26.95, 0.8642	27.11, 0.8523	25.30, 0.8411	27.75, 0.9103	27.68, 0.8731	31.29, 0.9291	**32.30, 0.9324**
stuffed_toys	32.07, 0.9260	28.56, 0.9180	29.43, 0.8735	33.74, 0.9457	32.86, 0.9284	**35.41, 0.9550**	35.04, 0.9548
paints	26.69, 0.8943	27.95, 0.9159	27.31, 0.9041	28.46, 0.9402	27.55, 0.9054	31.76, 0.9593	**32.53, 0.9621**
thread_spools	29.01, 0.8647	29.52, 0.8929	27.25, 0.8125	29.40, 0.8991	29.62, 0.8714	32.85, 0.9231	**34.07, 0.9235**
clay	34.48, 0.9150	26.51, 0.8142	36.57, 0.9191	**40.41, 0.9723**	35.40, 0.9227	40.28, 0.9578	33.06, 0.9003
photo_and_face	33.30, 0.9343	35.41, **0.9571**	30.84, 0.8688	35.11, 0.9523	33.75, 0.9332	37.56, 0.9545	**38.62**, 0.9543
chart_and_stuffed_toy	25.19, 0.8703	28.73, 0.9104	24.65, 0.8275	26.21, 0.9050	25.82, 0.8758	**29.82, 0.9236**	29.45, 0.9232
beads	22.08, 0.6929	22.69, 0.7466	20.78, 0.6026	22.94, 0.7621	22.54, 0.7111	24.70, 0.7994	**25.49, 0.8205**
fake_and_real_food	32.67, 0.9001	30.37, 0.8725	29.25, 0.8543	32.54, 0.9269	33.30, 0.9049	35.66, **0.9387**	**36.33**, 0.9304
fake_and_real_lemon_slices	29.83, 0.8476	33.21, **0.9274**	28.16, 0.8216	30.69, 0.9234	30.39, 0.8572	34.39, 0.9232	**35.93**, 0.9263
fake_and_real_peppers	34.05, 0.9397	30.12, 0.8817	30.74, 0.8993	33.97, 0.9544	34.74, 0.9458	36.05, **0.9564**	**37.11**, 0.9469
flowers	31.83, 0.8906	30.04, 0.8823	27.06, 0.8206	32.51, 0.9197	32.18, 0.8961	34.36, 0.9320	**34.94, 0.9383**
fake_and_real_beers	33.65, 0.9566	31.70, 0.9460	32.31, 0.9311	35.68, 0.9686	34.16, 0.9574	37.99, 0.9772	**38.29, 0.9785**
cloth	24.20, 0.6664	27.62, 0.7899	18.77, 0.4643	22.24, 0.6221	24.46, 0.6747	26.50, 0.7801	**29.54, 0.8451**
jelly_beans	24.54, 0.7913	27.10, 0.8539	23.21, 0.7481	24.79, 0.8459	25.19, 0.8114	27.44, 0.8830	**29.26, 0.9059**
sponges	29.54, 0.9298	22.76, 0.8136	30.70, 0.9349	33.70, 0.9699	30.29, 0.9349	**34.86, 0.9713**	31.19, 0.9549
watercolors	26.11, 0.8465	29.76, 0.9238	22.29, 0.7167	26.21, 0.8547	26.51, 0.8529	29.14, 0.9129	**31.67, 0.9391**
egyptian_statue	33.60, 0.9513	36.76, 0.9646	30.65, 0.9008	33.63, 0.9528	34.07, 0.9511	37.32, 0.9623	**39.35, 0.9682**
fake_and_real_sushi	33.29, 0.9507	34.47, 0.9443	30.85, 0.8926	33.26, 0.9685	33.76, 0.9513	35.95, **0.9689**	**37.04**, 0.9571
hairs	34.28, 0.9165	37.08, 0.9509	30.89, 0.8365	34.33, 0.9358	34.73, 0.9183	39.02, 0.9527	**40.83, 0.9627**
oil_painting	26.35, 0.6975	28.71, 0.8035	22.59, 0.6426	27.47, 0.7879	26.80, 0.7067	30.38, 0.8552	**32.12, 0.8785**
pompoms	30.79, 0.8692	25.61, 0.8385	27.21, 0.7740	31.32, 0.8823	31.56, 0.8757	**32.84**, 0.8966	31.40, **0.9110**
real_and_fake_peppers	34.55, 0.9421	31.68, 0.9014	32.29, 0.9187	35.17, 0.9645	35.11, 0.9454	38.78, **0.9727**	40.14, **0.9599**
fake_and_real_lemons	34.53, 0.9415	32.57, 0.9199	31.54, 0.9060	34.96, 0.9638	35.00, 0.9427	38.57, **0.9708**	40.03, 0.9601
glass_tiles	24.18, 0.7783	24.93, 0.8378	23.82, 0.7835	25.07, 0.8693	24.72, 0.7883	28.90, 0.9039	**30.17, 0.9210**
superballs	33.48, 0.8501	29.69, 0.8693	32.67, 0.8615	34.67, **0.9522**	34.17, 0.8605	**36.64**, 0.9164	34.34, 0.9180
fake_and_real_tomatoes	30.58, 0.9418	31.85, 0.9409	29.51, 0.8905	31.04, 0.9552	31.12, 0.9426	34.19, **0.9601**	**35.38**, 0.9526
cd	32.21, 0.8655	27.86, 0.9112	31.49, 0.8523	**35.70, 0.9497**	32.72, 0.8706	35.37, 0.9304	30.45, **0.9497**
fake_and_real_strawberries	31.59, 0.8855	33.60, 0.9089	28.90, 0.8542	31.69, 0.9320	32.21, 0.8925	35.76, 0.9348	**37.53, 0.9384**
face	34.06, 0.9493	33.87, 0.9590	31.55, 0.8749	34.99, 0.9564	34.51, 0.9475	38.32, 0.9648	**39.67, 0.9652**
balloons	35.42, 0.9648	30.48, 0.9296	34.00, 0.9503	37.05, 0.9757	35.58, 0.9640	**40.25, 0.9819**	38.37, 0.9780
real_and_fake_apples	37.34, 0.9625	36.32, 0.9201	34.17, 0.9071	37.99, **0.9703**	37.82, 0.9610	40.38, 0.9697	**41.54,** 0.9535

[*] Our methods performs 100 iteration. 2DTV performs 250 iteration beacause DeSCI relies on it as an initialization. 3DTV performs 100 iteration. FFDNet performs 100 iteration.
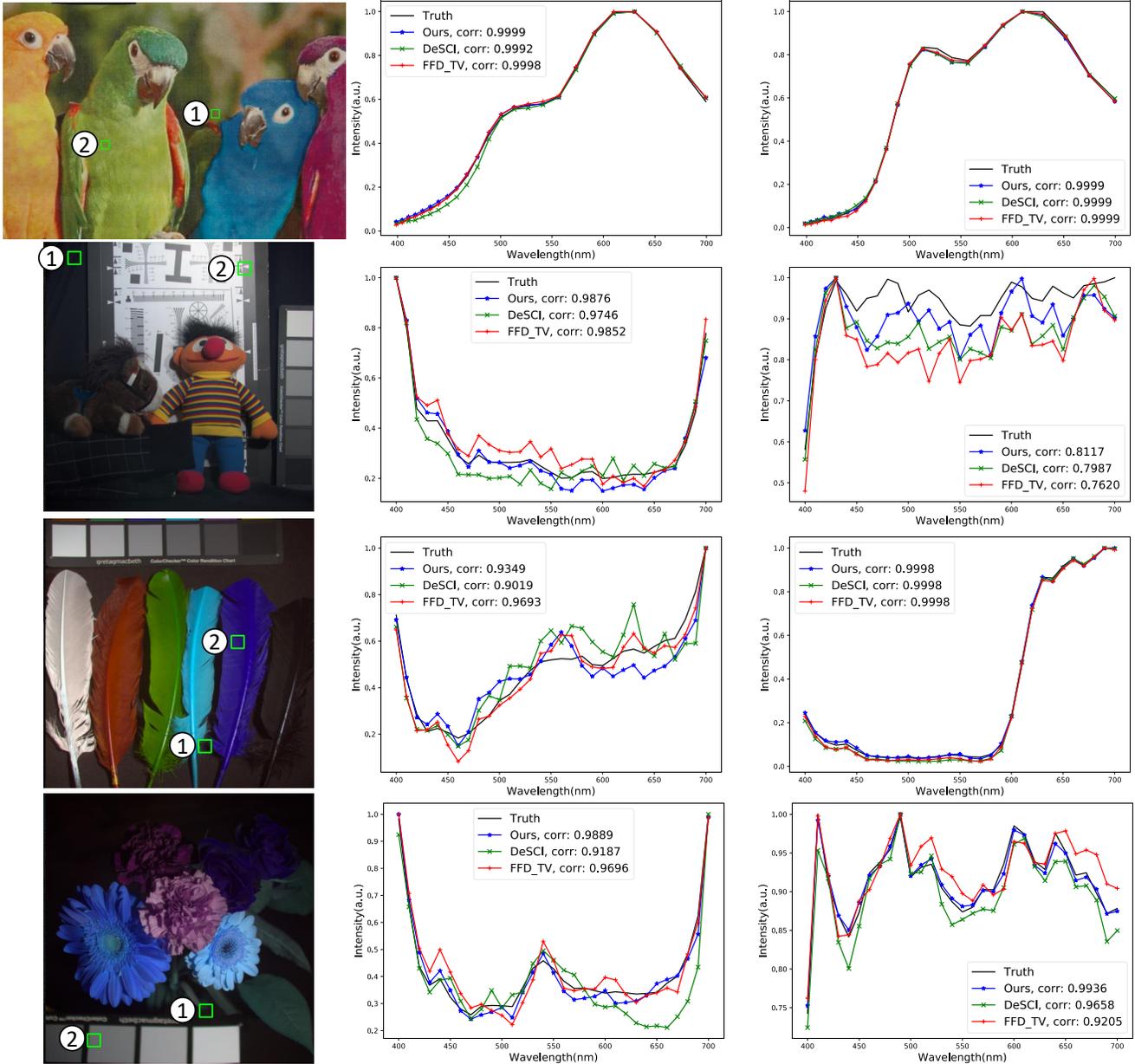
Figure 2. Simulated data: Spectral curves of reconstruction. We select `Bird`, `Toy`, feathers, and flowers hyperspectral image for comparison. The areas bounding by the green box are selected randomly.
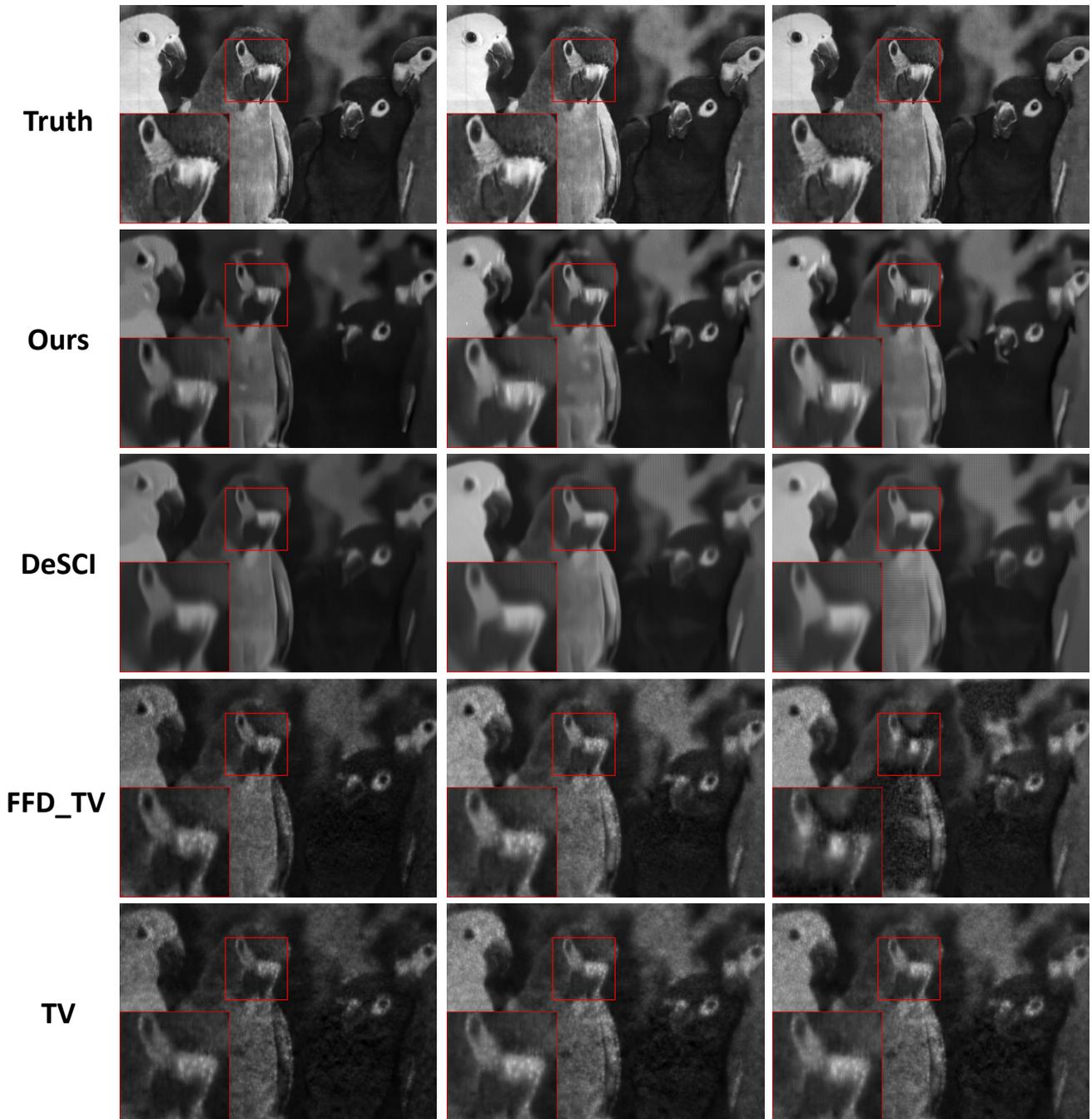
Figure 3. Real data: `Bird`. The three frames are at wavelengths 591.02nm, 630.13nm, and 651.74nm.